

Probability Theory, An Analytic View

The third edition of this highly regarded text provides a rigorous, yet entertaining, introduction to probability theory and the analytic ideas and tools on which the modern theory relies. The main changes are the inclusion of the Gaussian isoperimetric inequality plus many improvements and clarifications throughout the text. With more than 750 exercises, it is ideal for first-year graduate students with a good grasp of undergraduate probability theory and analysis.

Starting with results about independent random variables, the author introduces weak convergence of measures and its application to the central limit theorem, and infinitely divisible laws and their associated stochastic processes. These are followed by the introduction of conditional expectation values and martingales. The context then shifts to infinite dimensions, where Gaussian measures and weak convergence of measures are studied. The remainder is devoted to the mutually beneficial connection between probability theory and partial differential equations, culminating in an explanation of the relationship of Brownian motion to classical potential theory.

Daniel W. Stroock is Simons Professor Emeritus of Mathematics at the Massachusetts Institute of Technology. He has published numerous articles and books, most recently *Elements of Stochastic Calculus and Analysis* (2018) and *Gaussian Measures in Finite and Infinite Dimensions* (2023).

Probability Theory, An Analytic View

Third Edition

Daniel W. Stroock

Massachusetts Institute of Technology



Cambridge University Press & Assessment
978-1-009-54900-4 — Probability Theory, An Analytic View 3rd Edition
Daniel W. Stroock
Frontmatter
[More Information](#)



CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781009549004

DOI: 10.1017/9781009549035

© Daniel W. Stroock 1994, 2011, 2025

This publication is in copyright. Subject to statutory exception and to the provisions
of relevant collective licensing agreements, no reproduction of any part may take
place without the written permission of Cambridge University Press & Assessment.

When citing this work, please include a reference to the DOI 10.1017/9781009549035

First edition published 1994
First paperback edition 2000
Second edition published 2011
Third edition published 2025

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data

Names: Stroock, Daniel W., author.

Title: Probability theory, an analytic view / Daniel W. Stroock,
Massachusetts Institute of Technology.

Description: Third edition. | Cambridge, United Kingdom ; New York, NY, USA :
Cambridge University Press, [2025] | Includes bibliographical references and index.

Identifiers: LCCN 2024020566 (print) | LCCN 2024020567 (ebook) |
ISBN 9781009549004 (paperback) | ISBN 9781009549035 (ebook)

Subjects: LCSH: Probabilities.

Classification: LCC QA273 .S763 2025 (print) | LCC QA273 (ebook) | DDC
519.2—dc23/eng/20240715

LC record available at <https://lcn.loc.gov/2024020566>

LC ebook record available at <https://lcn.loc.gov/2024020567>

ISBN 978-1-009-54900-4 Paperback

Cambridge University Press & Assessment has no responsibility for the persistence
or accuracy of URLs for external or third-party internet websites referred to in this
publication and does not guarantee that any content on such websites is, or will
remain, accurate or appropriate.

**This book is dedicated to my teachers:
M. Kac, H. P. McKean, and S. R. S. Varadhan
and those who choose to read it.**

Contents

<i>Preface</i>	<i>page</i> xiii
<i>Notation</i>	xx
1 Sums of Independent Random Variables	1
1.1 Independence	1
1.1.1 Mutually Independent σ -Algebras	1
1.1.2 Mutually Independent Functions	3
1.1.3 The Rademacher Functions	4
1.1.4 Exercises for §1.1	6
1.2 The Weak Law of Large Numbers	12
1.2.1 Orthogonal Random Variables	13
1.2.2 Mutually Independent Random Variables	13
1.2.3 Approximate Identities	15
1.2.4 Exercises for §1.2	18
1.3 Cramér’s Theory of Large Deviations	19
1.3.1 Exercises for §1.3	27
1.4 The Strong Law of Large Numbers	31
1.4.1 Exercises for §1.4	37
1.5 The Law of the Iterated Logarithm	43
1.5.1 Exercises for §1.5	49
2 The Central Limit Theorem	51
2.1 The Basic Central Limit Theorem	52
2.1.1 Lindeberg’s Theorem	52
2.1.2 The Central Limit Theorem	54
2.1.3 Exercises for §2.1	56
2.2 The Berry–Esseen Theorem via Stein’s Method	62
2.2.1 L^1 -Berry–Esseen	63
2.2.2 The Classical Berry–Esseen Theorem	65
2.2.3 Exercises for §2.2	71
2.3 Some Extensions of the Central Limit Theorem	71
2.3.1 The Fourier Transform	71
2.3.2 Multidimensional Central Limit Theorem	73
2.3.3 Higher Moments	76
2.3.4 Exercises for §2.3	78
2.4 An Application to Hermite Multipliers	84
2.4.1 Hermite Multipliers	84
2.4.2 Beckner’s Theorem	88

viii	<i>Contents</i>	
	2.4.3 Applications of Beckner’s Theorem	92
	2.4.4 Exercises for §2.4	96
3	Infinitely Divisible Laws	101
3.1	Convergence of Measures on \mathbb{R}^N	101
3.1.1	Sequential Compactness in $\mathbf{M}_1(\mathbb{R}^N)$	102
3.1.2	Lévy’s Continuity Theorem	103
3.1.3	Exercises for §3.1	104
3.2	The Lévy–Khinchine Formula	107
3.2.1	$\mathcal{I}(\mathbb{R}^N)$ Is the Weak Closure of $\mathcal{P}(\mathbb{R}^N)$	108
3.2.2	The Formula	111
3.2.3	Exercises for §3.2	119
3.3	Stable Laws	121
3.3.1	General Results	121
3.3.2	α -Stable Laws	123
3.3.3	Exercises for §3.3	128
4	Lévy Processes	131
4.1	Stochastic Processes, Some Generalities	132
4.1.1	The Space $D(\mathbb{R}^N)$	133
4.1.2	Jump Functions	135
4.1.3	Exercises for §4.1	138
4.2	Discontinuous Lévy Processes	139
4.2.1	The Simple Poisson Process	140
4.2.2	Compound Poisson Processes	142
4.2.3	Poisson Jump Processes	147
4.2.4	Lévy Processes with Bounded Variation	148
4.2.5	General, Non-Gaussian Lévy Processes	150
4.2.6	Exercises for §4.2	152
4.3	Brownian Motion, the Gaussian Lévy Process	155
4.3.1	Deconstructing Brownian Motion	156
4.3.2	Lévy’s Construction of Brownian Motion	158
4.3.3	Kolmogorov’s Continuity Criterion	159
4.3.4	Brownian Paths Are Nondifferentiable	161
4.3.5	General Lévy Processes	163
4.3.6	Exercises for §4.3	164
5	Conditioning and Martingales	170
5.1	Conditioning	170
5.1.1	Kolmogorov’s Definition	171
5.1.2	Some Extensions	175
5.1.3	Exercises for §5.1	178
5.2	Discrete Parameter Martingales	181
5.2.1	Doob’s Inequality and Marcinkewitz’s Theorem	182
5.2.2	Doob’s Stopping Time Theorem	187
5.2.3	Martingale Convergence Theorem	188
5.2.4	Reversed Martingales and De Finetti’s Theory	191
5.2.5	Exercises for §5.2	195

6	Some Extensions and Applications of Martingale Theory	200
6.1	Some Extensions	200
6.1.1	Martingale Theory for a σ -Finite Measure Space	200
6.1.2	Banach Space-Valued Martingales	205
6.1.3	Exercises for §6.1	206
6.2	Burkholder’s Inequality	209
6.2.1	Burkholder’s Comparison Theorem	209
6.2.2	Burkholder’s Inequality	213
6.2.3	Exercises for §6.3	214
6.3	Elements of Ergodic Theory	216
6.3.1	The Maximal Ergodic Lemma	217
6.3.2	Birkhoff’s Ergodic Theorem	220
6.3.3	Ergodic Decomposition	223
6.3.4	Unique Ergodicity	226
6.3.5	Stationary Sequences	229
6.3.6	Continuous Parameter Ergodic Theory	231
6.3.7	Exercises for §6.3	233
7	Continuous Parameter Martingales	235
7.1	Continuous Parameter Martingales	235
7.1.1	Progressively Measurable Functions	235
7.1.2	Martingales: Definition and Examples	236
7.1.3	Basic Results	239
7.1.4	Stopping Times and Stopping Theorems	240
7.1.5	An Integration by Parts Formula	244
7.1.6	Exercises for §7.1	247
7.2	Brownian Motion and Martingales	249
7.2.1	Lévy’s Characterization of Brownian Motion	249
7.3	Doob–Meyer Decomposition, an Easy Case	251
7.3.1	Burkholder’s Inequality Again	255
7.3.2	Exercises for §7.2 and §7.3	256
7.4	The Reflection Principle Revisited	258
7.4.1	Reflecting Symmetric Lévy Processes	258
7.4.2	Reflected Brownian Motion	259
7.4.3	Exercises for §7.4	262
8	Gaussian Measures on a Banach Space	264
8.1	The Classical Wiener Space	264
8.1.1	Classical Wiener Measure	264
8.1.2	The Classical Cameron–Martin Space	267
8.1.3	Exercises for §8.1	270
8.2	A Structure Theorem for Gaussian Measures	270
8.2.1	Fernique’s Theorem	270
8.2.2	The Basic Structure Theorem	272
8.2.3	The Cameron–Martin Space	275
8.2.4	Exercises for §8.2	278
8.2.5	The Ornstein–Uhlenbeck Process	281
8.2.6	Ornstein–Uhlenbeck as an Abstract Wiener Space	283
8.3	Wiener’s Construction of Abstract Wiener Space	285
8.3.1	Wiener Series	285

x	<i>Contents</i>	
	8.3.2 Pinned Brownian Motion	288
	8.3.3 Orthogonal Invariance	290
	8.3.4 Exercises for §8.3	293
8.4	The Gaussian Isoperimetric Inequality	297
	8.4.1 Strassen’s Law of the Iterated Logarithm	300
	8.4.2 Exercises for §8.4	303
9	Convergence of Measures on a Polish Space	304
9.1	Prohorov–Varadarajan Theory	304
	9.1.1 Some Background	304
	9.1.2 The Weak Topology	306
	9.1.3 The Lévy Metric and Completeness of $\mathbf{M}_1(E)$	312
	9.1.4 Kolmogorov’s Consistency Theorem	316
	9.1.5 Exercises for §9.1	318
9.2	Regular Conditional Probability Distributions	321
	9.2.1 Fibering a Measure	323
	9.2.2 Representing Lévy Measures via the Itô Map	324
	9.2.3 Exercises for §9.2	326
9.3	Donsker’s Invariance Principle	326
	9.3.1 Donsker’s Theorem	327
	9.3.2 Rayleigh’s Random Flights Model	330
	9.3.3 Exercises for §9.3	332
10	Wiener Measure and Partial Differential Equations	334
10.1	Martingales and Partial Differential Equations	334
	10.1.1 Localizing and Extending Martingale Representations	335
	10.1.2 The Minimum Principles	337
	10.1.3 The Hermite Heat Equation	339
	10.1.4 The Arcsine Law	340
	10.1.5 Recurrence and Transience of Brownian Motion	344
	10.1.6 Exercises for §10.1	347
10.2	The Markov Property and Potential Theory	348
	10.2.1 The Markov Property for Wiener Measure	348
	10.2.2 Recurrence in One and Two Dimensions	349
	10.2.3 The Dirichlet Problem	350
	10.2.4 Exercises for §10.2	356
10.3	Other Heat Kernels	359
	10.3.1 A General Construction	360
	10.3.2 The Dirichlet Heat Kernel	361
	10.3.3 Feynman–Kac Heat Kernels	365
	10.3.4 Ground States and Associated Measures on Pathspace	368
	10.3.5 Producing Ground States	374
	10.3.6 Exercises for §10.3	378
11	Some Classical Potential Theory	384
11.1	Uniqueness Refined	384
	11.1.1 The Dirichlet Heat Kernel Again	384
	11.1.2 Exiting through $\partial_{\text{reg}}G$	387
	11.1.3 Applications to Questions of Uniqueness	390
	11.1.4 Harmonic Measure	395

	<i>Contents</i>	xi
11.1.5	Exercises for §11.1	399
11.2	The Poisson Problem and Green’s Functions	401
11.2.1	Green’s Functions when $N \geq 3$	402
11.2.2	Green’s Functions when $N \in \{1, 2\}$	403
11.2.3	Exercises for §11.2	411
11.3	Excessive Functions, Potentials, and Riesz Decompositions	412
11.3.1	Excessive Functions	412
11.3.2	Potentials and Riesz Decomposition	413
11.3.3	Exercises for §11.3	419
11.4	Capacity	420
11.4.1	The Capacitory Potential	420
11.4.2	The Capacitory Distribution	423
11.4.3	Wiener’s Test	426
11.4.4	Asymptotic Expressions Involving Capacity	429
11.4.5	Exercises for §11.4	435
	<i>References</i>	437
	<i>Index</i>	440

Preface

Because they have been the most read parts of previous editions, I have chosen to include excerpts from their prefaces.

From the Preface to the First Edition

When writing a graduate-level mathematics book during the last decade of the twentieth century, one probably ought not to inquire too closely into one's motivation. In fact, if one's own pleasure from the exercise is not sufficient to justify the effort, then one should seriously consider dropping the project. Thus, to those who (either before or shortly after opening it) ask *for whom was this book written*, my pale answer is *me*; and, for this reason, I thought that I should preface this preface with an explanation of who I am and what were the peculiar educational circumstances that eventually gave rise to this somewhat peculiar book.

My own introduction to probability theory began with a private lecture from H. P. McKean, Jr. At the time, I was a (more accurately, *the*) graduate student of mathematics at what was then called The Rockefeller Institute for Biological Sciences. My official mentor there was M. Kac, whom I had cajoled into becoming my adviser after a year during which I had failed to insert even one microelectrode into the optic nerve of innumerable limuli. However, as I soon came to realize, Kac had accepted his role on the condition that it would not become a burden. In particular, he had no intention of wasting much of his own time on a reject from the neurophysiology department. On the other hand, he was most generous with the time of his younger associates, and that is how I wound up in McKean's office. Never one to bore his listeners with a lot of dull preliminaries, McKean launched right into a wonderfully lucid explanation of P. Lévy's interpretation of the infinitely divisible laws. I have to admit that my appreciation of the lucidity of his lecture arrived nearly a decade after its delivery, and I can only hope that my reader will reserve judgment of my own presentation for an equal length of time.

In spite of my perplexed state at the end of McKean's lecture, I was sufficiently intrigued to delve into the readings that he suggested at its conclusion. Knowing that the only formal mathematics courses that I would be taking during my graduate studies would be given at New York University (NYU) and guessing that those courses would be oriented toward the analysis of partial differential equations, McKean directed me to material that would help me understand the connections between partial differential equations and probability theory. In particular, he suggested that I start with the, then recently translated, two articles by E. B. Dynkin that had appeared originally in the famous 1956 volume of *Teoriya Veroyatnostei i ee Primeneniya*. Dynkin's articles turned out to be a godsend. They were beautifully crafted

to tell the reader enough so that he or she could understand the ideas and not so much that he or she would become bored by them. In addition, they gave me an introduction to a host of ideas and techniques (e.g., stopping times and the strong Markov property), all of which Kac himself consigned to the category of overelaborated measure theory. In fact, it would be reasonable to say that my thesis was simply the application of techniques that I picked up from Dynkin to a problem that I picked up by reading some notes by Kac. Of course, along the way, I profited immeasurably from continued contact with McKean, a large number of courses at NYU (particularly ones taught by M. Donsker, F. John, and L. Nirenberg), and my increasingly animated conversations with S. R. S. Varadhan.

As I trust the preceding description makes clear, my graduate education was anything but deprived; I had ready access to some of the very best analysts of the day. On the other hand, I never had a *proper* introduction to my field, probability theory. The first time that I ever summed independent random variables was when I was summing them in front of a class at NYU. Thus, although I now admire the magnificent body of mathematics created by A. N. Kolmogorov, P. Lévy, and the other twentieth-century heroes of the field, I am not a *dyed-in-the-wool* probabilist (i.e., what Donsker would have called a true *coin tosser*). In particular, I have never been able to develop sufficient sensitivity to the distinction between a *proof* and a *probabilistic proof*. To me, a proof is clearly *probabilistic* only if its punch line comes down to an argument like $\mathbb{P}(A) \leq \mathbb{P}(B)$ because $A \subseteq B$; and there are breathtaking examples of such arguments. However, to base an entire book on these examples would require a level of genius that I do not possess. In fact, I myself enjoy probability theory best when it is inextricably interwoven with other branches of mathematics and not when it is presented as an entity unto itself. For this reason, the reader should not be surprised to find some of the material presented in this book *does not belong here*; but I hope that they will make an effort to figure out why I disagree with them.

Preface to the Second Edition

My favorite “preface to a second edition” is the one which G. N. Watson wrote for the second edition of his famous treatise on Bessel functions. The first edition appeared in 1922, the second came out in 1941, and Watson had originally intended to stay abreast of developments and report on them in the second edition. However, in his preface to the second edition, Watson admits that his interest in the topic has “waned” during the intervening years and apologizes that, as a consequence, the new edition contains less new material than he had once thought it would.

My excuse for not incorporating more new material into this second edition is related to but somewhat different from Watson’s. In my case, what has waned is not my interest in probability theory but instead my ability to assimilate the transformations that the subject has undergone. When I was a student, probabilists were still working out the ramifications of Kolmogorov’s profound insights into the connections between probability and analysis, and I have spent my career investigating and exploiting those connections. However, about the time when the first edition of this book was published, probability theory began a return to its origins in combinatorics, a topic in which my abilities are woefully deficient. Thus, although I suspect that, for at least a decade, the most exciting developments in the field will have a strong combinatorial component, I have not attempted to prepare my readers for those

developments. My decision not to incorporate more combinatorics into this new edition in no way reflects my assessment of the direction in which probability is likely to go. Instead, it reflects my assessment of my own inability to do justice to the beautiful combinatorial ideas that have been introduced in the recent past.

In spite of the preceding admission, I believe that the material in this book remains valuable and that, no matter how probability theory evolves, the ideas and techniques presented here will continue to play an important role. Furthermore, I have made some substantive changes. In particular, I have given more space to infinitely divisible laws and their associated Lévy processes, both of which are now developed in \mathbb{R}^N rather than just in \mathbb{R} . In addition, I have added an entire chapter devoted to Gaussian measures in infinite dimensions from the perspective of the Segal–Gross school. Not only have recent developments in Malliavin calculus and conformal field theory sparked renewed interest in this topic, but it seems to me that most modern texts pay either no or too little attention to this beautiful material. Missing from the new edition is the treatment of singular integrals. I included it in the first edition in the hope that it would elucidate the similarity between cancellations that underlie martingale theory, especially Burkholder’s inequality, and Calderon–Zygmund theory. I still believe that these similarities are worth thinking about, but I have decided that my explanation of them led me too far astray and was more of a distraction than a pedagogically valuable addition.

Besides those mentioned earlier, minor changes have been made throughout. For one thing, I have spent a lot of time correcting old errors and, undoubtedly, introducing new ones. Second, I have made several organizational changes and others of which are remedial.

Preface to the Third Edition

Watson had the good sense not to publish a third edition of his book, but, had he chosen to do so, I am sure that his reservations would have been even greater than those he expressed about publishing his second edition. My decision to publish a third edition was motivated in part by the hope that its contents might cause indigestion in the memory bank of an AI system, and in part by the desire to correct the errors in an ever-growing list that I and others have found in the second edition. In addition to correcting those errors, I wanted to make a few more substantive changes. The most important of these are in Chapter 8, where I have replaced some of the general theory of abstract Wiener spaces with a proof of the Gaussian isoperimetric inequality. Elsewhere, I have made some changes that I hope will clarify the presentation, and I have included a few new exercises.

Summary

1: Chapter 1 contains a sampling of the standard, point-wise convergence theorems dealing with partial sums of independent random variables. These include the Weak and Strong Laws of Large Numbers as well as Hartman–Wintner’s Law of the Iterated Logarithm. In preparation for the law of the iterated logarithm, Cramér’s theory of large deviations from the law of large numbers is developed in §1.3. Everything here is very standard, although I

feel that my passage from the bounded to the general case of the law of the iterated logarithm has been considerably smoothed by the ideas that I learned in conversation with M. Ledoux.

2: The whole of Chapter 2 is devoted to the classical Central Limit Theorem. After an initial (and slightly flawed) derivation of the basic result via moment considerations, Lindeberg's general version is derived in §2.1. Although Lindeberg's result has become a *sine qua non* in the writing of probability texts, the Berry–Esseen estimate has not. Indeed, until recently, the Berry–Esseen estimate required a good many somewhat tedious calculations with characteristic functions (i.e., Fourier transforms), and most recent authors seem to have decided that the rewards did not justify the effort. I was inclined to agree with them until P. Diaconis brought to my attention E. Bolthausen's adaptation of C. Stein's techniques (the so-called *Stein's method*) to give a proof that is not only brief but also, to me, aesthetically pleasing. In any case, no use of Fourier methods is made in the derivation given in §2.2. On the other hand, Fourier techniques are introduced in §2.3, where it is shown that even elementary Fourier analytic tools lead to important extensions of the basic Central Limit Theorem to more than one dimension. Finally, in §2.4, the Central Limit Theorem is applied to the study of Hermite multipliers and (following Wm. Beckner) is used to derive both E. Nelson's hypercontraction estimate for the Mehler kernel and Beckner's own estimate for the Fourier transform. I am afraid that, with this flagrant example of *the sort of thing that does not belong here*, I may be trying the patience of my purist colleagues. However, I hope that their indignation will be somewhat assuaged by the fact that rest of the book is essentially independent of the material in §2.4.

3: This chapter is devoted to the study of infinitely divisible laws. It begins in §3.1 with a few refinements (especially the Lévy Continuity Theorem) of the Fourier techniques introduced in §2.3. These play a role in §3.2, where the Lévy–Khinchine formula is first derived and then applied to the analysis of stable laws.

4: In Chapter 4, I construct the Lévy processes (a.k.a. independent increment processes) corresponding to infinitely divisible laws. Section 4.1 provides the requisite information about the pathspace $D(\mathbb{R}^N)$ of right-continuous paths with left limits, and §4.2 gives the construction of Lévy processes with discontinuous paths, the ones corresponding to infinitely divisible laws having no Gaussian part. Finally, in §4.3, I construct Brownian motion, the Lévy process with continuous paths, following the prescription given by Lévy. This section also contains a derivation of Kolmogorov's continuity criterion for general Banach space-valued stochastic processes.

5: Because they are not needed earlier, conditional expectations do not appear until Chapter 5. The advantage gained by this postponement is that, by the time I introduce them, I have an ample supply of examples to which conditioning can be applied; the disadvantage is that, with considerable justice, many probabilists feel that one is not doing *probability theory* until one is conditioning. Be that as it may, Kolmogorov's definition is given in §5.1 and is shown to extend naturally to both σ -finite measure spaces and random variables with values in a Banach space. Section 5.2 presents Doob's basic theory of real-valued, discrete parameter martingales: Doob's Inequality, his Stopping Time Theorem, and his Martingale

Convergence Theorem. In the last part of §5.2, I introduce reversed martingales and apply them to De Finetti's theory of exchangeable random variables.

6: Chapter 6 opens with extensions of martingale theory in two directions: to σ -finite measures and to random variables with values in a Banach space. In §6.2, I prove Burkholder's Inequality for martingales with values in a Hilbert space. The derivation that I give is essentially the same as Burkholder's second proof, the one that gives optimal constants. Finally, the results in §6.1 are used in §6.3 to derive Birkhoff's Individual Ergodic Theorem and a couple of its applications.

7: Section 7.1 provides a brief introduction to the theory of martingales with a continuous parameter. As anyone at all familiar with the topic knows, anything approaching a full account of this theory requires much more space than a book like this can provide. Thus, I deal with only its most rudimentary aspects, which, fortunately, are sufficient for the applications to Brownian motion that I have in mind. Namely, in §7.2, I first discuss the intimate relationship between continuous martingales and Brownian motion (Lévy's martingale characterization of Brownian motion), then derive the simplest (and perhaps most widely applied) case of the Doob–Meyer Decomposition Theory, and finally show what Burkholder's Inequality looks like for continuous martingales. In the concluding section, §7.3, the results in §7.1 and §7.2 are applied to derive the Reflection Principle for Brownian motion.

8: In §8.1, I formulate the description of Brownian motion in terms of its Gaussian, as opposed to its independent increment, properties. More precisely, following I. M. Segal and L. Gross, I attempt to convince the reader that Wiener measure (i.e., the distribution of Brownian motion) would like to be the standard Gauss measure on the Hilbert space $\mathbf{H}(\mathbb{R}^N)$ of absolutely continuous paths that are 0 at $\mathbf{0}$ and possessing square integrable derivative, but, for technical reasons, cannot live there and has to settle for a Banach space in which $\mathbf{H}(\mathbb{R}^N)$ is densely embedded. Using Wiener measure as the model, in §8.2, I show that, at an abstract level, any Gaussian measure on an infinite-dimensional, separable Banach space shares the same structure as Wiener measure in the sense that there is always a densely embedded Hilbert space, known as the Cameron–Martin space, for which it would like to be the standard Gaussian measure but on which it does not fit. In order to carry out this program, I need to first prove X. Fernique's remarkable theorem for Gaussian measures on a Banach space. In §8.3, I begin by going in the opposite direction, showing how to pass from a Hilbert space H to a Gaussian measure on a Banach space E for which H is the Cameron–Martin space. The rest of §8.3 gives two applications: one to “pinned Brownian” motion and the other to a very general statement of orthogonal invariance for Gaussian measures. The chapter concludes with a derivation in §8.4 of Schilder's Law of the Iterated Logarithm for Brownian motion.

9: The central topic here is the abstract theory of weak convergence of probability measures on a Polish space. The basic theory is developed in §9.1. In §9.2, I apply the theory to prove the existence of regular conditional probability distributions, and in §9.3, I use it to derive Donsker's Invariance Principle (i.e., the pathspace statement of the Central Limit Theorem).

10: Chapter 10 is an introduction to the connections between probability theory and partial differential equations. At the beginning of §10.1, I show that martingales provide a link between probability theory and partial differential equations. More precisely, I show how to represent in terms of Wiener integrals solutions to parabolic and elliptic partial differential equations in which the Laplacian is the principle part. In the second part of §10.1, I derive the Feynman–Kac formula and use it to calculate various Wiener integrals. In §10.2, I introduce the Markov property of Wiener measure and show how it not only allows one to evaluate other Wiener integrals in terms of solutions to elliptic partial differential equations but also enables one to prove interesting facts about solutions to such equations as a consequence of their representation in terms of Wiener integrals. Continuing in the same spirit, I show in §10.2 how to represent solutions to the Dirichlet problem in terms of Wiener integrals, and in §10.3, I use Wiener measure to construct and discuss heat kernels related to the Laplacian and discuss ground states (a.k.a. stationary measures) for them.

11: The final chapter is an extended example of the way in which probability theory meshes with other branches of analysis, and the example that I have chosen is the marriage between Brownian motion and classical potential theory. Like an ideal marriage, this one is simultaneously intimate and mutually beneficial to both partners. Indeed, the more one knows about it, the more convinced one becomes that the properties of Brownian paths are a perfect reflection of properties of harmonic functions, and vice versa. In any case, in §11.1, I sharpen the results in §10.2.3 and show that, in complete generality, the solution to the Dirichlet problem is given by the Wiener integral of the boundary data evaluated at the place where Brownian paths exit from the region. Next, in §11.2, I discuss the Green’s function for a region and explain how its existence reflects the recurrence and transience properties of Brownian paths. In preparation for §11.4, §11.3 is devoted to the Riesz Decomposition Theorem for excessive functions. Finally, in §11.4, I discuss the capacity of regions, derive Chung’s representation of the capacitory measure in terms of the last place where a Brownian path visits a region, apply the probabilistic interpretation of capacity to give a derivation of Wiener’s test for regularity, and conclude with two asymptotic calculations, one by F. Spitzer and the other by G. Hunt, in which capacity plays a crucial role.

Suggestions about the Use of This Book

In spite of the realistic assessment contained in the first paragraph of its preface, when I wrote the first edition of this book, I harbored the naïve hope that it might become *the standard* graduate text in probability theory. By the time that I started preparing the second edition, I was significantly older and far less naïve about its prospects, and by now, I am completely resigned to its limited circulation. Although the earlier editions have their admirers, they have done little to damage the sales record of their competitors. In particular, this book has seldom been adopted as the text for courses in probability, and I doubt that the appearance of a new edition will improve its rate of adoption. Nonetheless, I close this preface with a few suggestions for anyone who does choose to base a course on it.

I am well aware that, except for those who find their way into the poorly stocked library of some prison camp, few copies of this book will be read from cover to cover. For this reason, I have attempted to organize it in such a way that, with the help of the table of contents

and the index, a reader can select a path that does not require their reading all the sections preceding the information that they seek. For example, the contents of §1.1, §1.2, §1.4, §2.1, §2.3, and Chapter 5 constitute the backbone of a one-semester, graduate-level introduction to probability theory. What one attaches to this backbone depends on the speed with which these sections are covered and the content of the courses is the introduction. If the goal is to prepare the students for a career as a “quant” in what is left of the financial industry, an obvious choice is §4.3 and as much of Chapter 7 as time permits, thereby giving one’s students a reasonably solid introduction to Brownian motion. On the other hand, if one wants the students to appreciate that white noise is not the only noise that they may encounter in life, one might defer the discussion of Brownian motion and replace it with the material in Chapter 3 and §4.1 and §4.2.

Alternatively, one might use this book in a more advanced course. An introduction to stochastic processes with an emphasis on their relationship to partial differential equations can be constructed out of Chapters 4, 7, 10, and 11, and §4.3 combined with Chapter 7 could be used to provide background for a course on Gaussian processes.

Whatever route one takes through this book, it will be a great help to your students for you to suggest that they consult other texts. Indeed, it is a familiar fact that the third book one reads on a subject is always the most lucid, and so one should suggest at least two other books. Among the many excellent choices available, I mention: Wm. Feller’s *An Introduction to Probability Theory and Its Applications, Vol. II* [20] and M. Loève’s classic *Probability Theory* [41] (now published by Dover Press), although both are somewhat dated. In addition, for background, precision (including accuracy of attribution), and supplementary material, R. Dudley’s *Real Analysis and Probability* [18] is superb, and R. Durrett’s *Probability: Theory and Examples* [19] is deservedly popular.

Notation

General

\mathbb{R} & \mathbb{C}	The real line and complex plane
\mathbb{N}	The nonnegative integers
\mathbb{Z} & \mathbb{Z}^+	The integers and the positive integers
\mathbb{Q}	The rational numbers
a^+ & a^-	The positive and negative parts of $a \in \mathbb{R}$
$a \vee b$ & $a \wedge b$	The maximum and minimum of $a \in \mathbb{R}$ and $b \in \mathbb{R}$
$f \upharpoonright S$	The restriction of the function f to the set S
$\ \cdot\ _{\mathbf{u}}$	The uniform norm of a function
$\ \psi\ _{[a,b]}$	The supremum of a path $\psi \upharpoonright [a,b]$
$\text{var}_{[a,b]}(\psi)$	The variation norm of a path $\psi \upharpoonright [a,b]$
\mathbb{S}^{N-1} & ω_{N-1}	The unit sphere in \mathbb{R}^N and its surface area
$[t]$	The integer part of $t \in \mathbb{R}$
$\lceil t \rceil$	The smallest integer larger than or equal to $t \in \mathbb{R}$
$A^{\mathbb{C}}$	The complement of a set A
$B(a,r)$	The open ball centered at a of radius r
$K \subset\subset E$	To be read: “ K is a compact subset of E ”

Spaces

$C(E;F)$ & $C_{\mathbf{b}}(E;F)$	The continuous and bounded continuous functions mapping E to F
$C_{\mathbf{c}}(E;\mathbb{R})$	The compactly supported continuous functions from E into \mathbb{R}
$C^{\infty}(G;\mathbb{R})$ & $C^{\infty}(G;\mathbb{C})$	Infinitely \mathbb{R} - or \mathbb{C} -valued functions on $G \subseteq \mathbb{R}^N$
$C_{\mathbf{c}}^{\infty}(G;\mathbb{R})$ & $C_{\mathbf{c}}^{\infty}(G;\mathbb{C})$	Compactly supported functions in $C^{\infty}(G;\mathbb{R})$ or $C^{\infty}(G;\mathbb{C})$
$C^{1,2}(G;\mathbb{R})$	The \mathbb{R} - or \mathbb{C} -valued functions $f(t,x)$ on $G \subseteq \mathbb{R} \times \mathbb{R}^N$, which are continuously differentiable once in t and twice in x
$\mathcal{S}(\mathbb{R}^N;\mathbb{R})$ & $\mathcal{S}(\mathbb{R}^N;\mathbb{C})$	The Schwartz test function space of C^{∞} , \mathbb{R} - or \mathbb{C} -valued functions with rapidly decreasing derivatives of all orders
$C(\mathbb{R}^N)$	The space $C([0,\infty);\mathbb{R}^N)$ of continuous, \mathbb{R}^N -valued paths

$D(\mathbb{R}^N)$	The space of right continuous, \mathbb{R}^N -valued paths on $[0, \infty)$ with left limits at each $t \in (0, \infty)$
$\mathbf{M}(E)$ & $\mathbf{M}_1(E)$	The finite and the probability measures on E
$L^p(\mu; \mathbb{R})$ & $L^p(\mu; \mathbb{C})$	The Lebesgue spaces L^p for the measure μ
$\mathcal{E}(G)$	The space of excessive functions on G

Measure Theoretic

\mathcal{B}_E	The Borel σ -algebra (field) for a topological space E
(E, \mathcal{F})	The measurable space E with σ -algebra \mathcal{F}
$B(E; \mathbb{R})$ & $B(E; \mathbb{C})$	The bounded, measurable \mathbb{R} - and \mathbb{C} -valued functions on E
$\mathbb{E}^\mu[X]$ & $\mathbb{E}^\mu[X, A]$	The expected value of X and $X \upharpoonright A$ with respect to the measure μ
$\mathbb{E}^\mu[X \Sigma]$	The conditional expectation value of X given the σ -algebra Σ
$\langle \varphi, \mu \rangle$	The integral of the function φ with respect to the measure μ
\hat{f} & \check{f}	The Fourier and inverse Fourier transforms of f
$f * g$	The convolution of the functions f and g
$\hat{\mu}$	The Fourier transform or characteristic function of the measure μ
$\mu * \nu$	The convolution of the measures μ and ν
$\mu \ll \nu$	The measure μ is absolutely continuous with respect to the measure ν
$\mu \perp \nu$	The measure μ is singular to the measure ν
$\mu_n \implies \mu$	The sequence $\{\mu: n \geq 1\}$ of measures converges weakly to the measure μ
$\Phi_*\mu$	The pushforward of the measure μ under the map Φ
λ_E	Lebesgue measure on $E \subseteq \mathbb{R}^N$
δ_a	The unit point measure at a
$\gamma_{\mathbf{m}, \mathbf{C}}$	The Gauss measure or normal distribution with mean \mathbf{m} and covariance \mathbf{C}
$N(\mathbf{m}, \mathbf{C})$	The set of random variables with distribution $\gamma_{\mathbf{m}, \mathbf{C}}$
$\sigma(\{X_i: i \in I\})$	The smallest σ -algebra with respect to which the random variables $\{X_i: i \in I\}$ are measurable
$\bigvee_{i \in I} \mathcal{F}_i$	The smallest σ -algebra containing $\bigcup_{i \in I} \mathcal{F}_i$

Functions and Operators

$g^{(N)}(t, \mathbf{x})$	The density of $\gamma_{0, \mathbf{I}}$
$p^G(t, \mathbf{x}, \mathbf{y})$	The transition probability function for Brownian motion killed at ∂G
\mathbf{P}_t^G	The operator with kernel $p^G(t, \mathbf{x}, \mathbf{y})$
$g^G(\mathbf{x}, \mathbf{y})$	The Green's function for the Dirichlet problem on G
\mathbf{G}^G	The operator with kernel $g^G(\mathbf{x}, \mathbf{y})$

Wiener Space

$\Theta(\mathbb{R}^N)$	The space of $\psi \in C(\mathbb{R}^N)$ such that $\lim_{t \rightarrow \infty} \frac{ \psi(t) }{1+t} = 0$
$\mathcal{W}^{(N)}$ & $\mathcal{W}_{\mathbf{x}}^{(N)}$	Wiener measure on $\Theta(\mathbb{R}^N)$ and Wiener measure on $\{\psi \in C(\mathbb{R}^N) : \psi(0) = \mathbf{x}\}$
(H, E, \mathcal{W}_H)	Abstract Wiener space with Cameron–Martin space H
Σ_s	The time shift map on $C(\mathbb{R}^N)$
δ_s	The differential time shift map Σ_s – ident on $C(\mathbb{R}^N)$