

## *A User's Guide to Measure Theoretic Probability*

This book grew from a need to teach a rigorous probability course to a mixed audience – statisticians, mathematically inclined biostatisticians, mathematicians, economists, and students of finance – at the advanced undergraduate/introductory graduate level, without the luxury of a course in measure theory as a prerequisite. The core of the book covers the basic topics of independence, conditioning, martingales, convergence in distribution, and Fourier transforms. In addition, there are numerous sections treating topics traditionally thought of as more advanced, such as coupling and the KMT strong approximation, option pricing via the equivalent martingale measure, and Fernique's inequality for Gaussian processes.

In a further break with tradition, the necessary measure theory is developed via the identification of integrals with linear functionals on spaces of measurable functions, allowing quicker access to the full power of the measure theoretic methods.

The book is not just a presentation of mathematically rigorous theory; it is also a discussion of why some of that theory takes its current form and how anyone could have thought of those clever ideas in the first place. It is intended as a secure starting point for anyone who needs to invoke rigorous probabilistic arguments and to understand what they mean.

*David Pollard* is Professor of Statistics and Mathematics at Yale University in New Haven, Connecticut. His interests center on probability, measure theory, theoretical and applied statistics, and econometrics. He believes strongly that research and teaching (at all levels) should be intertwined. His book, *Convergence of Stochastic Processes* (Springer-Verlag, 1984), successfully introduced many researchers and graduate students to empirical process theory.

CAMBRIDGE SERIES IN STATISTICAL AND PROBABILISTIC MATHEMATICS

---

*Editorial Board*

- R. Gill (Department of Mathematics, Utrecht University)  
B. D. Ripley (Department of Statistics, University of Oxford)  
S. Ross (Department of Industrial Engineering, University of California, Berkeley)  
M. Stein (Department of Statistics, University of Chicago)  
D. Williams (School of Mathematical Sciences, University of Bath)

This series of high quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

*Already Published*

1. *Bootstrap Methods and Their Application*, by A. C. Davison and D. V. Hinkley
2. *Markov Chains*, by J. Norris
3. *Asymptotic Statistics*, by A. W. van der Vaart
4. *Wavelet Methods for Time Series Analysis*, by Donald B. Percival and Andrew T. Walden
5. *Bayesian Methods: An Analysis for Statisticians and Interdisciplinary Researchers*, by Thomas Leonard and John S. J. Hsu
6. *Empirical Processes in M-Estimation*, by Sara van de Geer
7. *Numerical Methods of Statistics*, by John F. Monahan



DAVID POLLARD  
*Yale University*

---

*A User's Guide to  
Measure Theoretic  
Probability*



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](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521002899](http://www.cambridge.org/9780521002899)

© David Pollard 2002

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.

First published 2002

7th printing 2010

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

*Library of Congress Cataloging-in-Publication data*

Pollard, David, 1950–

A user's guide to measure theoretic probability / David Pollard.

p. cm. – (Cambridge series in statistical and probabilistic mathematics)

Includes bibliographical references and index.

ISBN 0-521-80242-3 – ISBN 0-521-00289-3 (pbk.)

1. Probabilities. 2. Measure theory. I. Title.

II. Cambridge series in statistical and probabilistic mathematics.

QA273 .P7735 2001

519.2 – dc21 2001035270

ISBN 978-0-521-80242-0 Hardback

ISBN 978-0-521-00289-9 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.

# Contents

PREFACE	xi
CHAPTER 1: MOTIVATION	
§1	Why bother with measure theory? 1
§2	The cost and benefit of rigor 3
§3	Where to start: probabilities or expectations? 5
§4	The de Finetti notation 7
*§5	Fair prices 11
§6	Problems 13
§7	Notes 14
CHAPTER 2: A MODICUM OF MEASURE THEORY	
§1	Measures and sigma-fields 17
§2	Measurable functions 22
§3	Integrals 26
*§4	Construction of integrals from measures 29
§5	Limit theorems 31
§6	Negligible sets 33
*§7	$L^p$ spaces 36
*§8	Uniform integrability 37
§9	Image measures and distributions 39
§10	Generating classes of sets 41
*§11	Generating classes of functions 43
§12	Problems 45
§13	Notes 51
CHAPTER 3: DENSITIES AND DERIVATIVES	
§1	Densities and absolute continuity 53
*§2	The Lebesgue decomposition 58
§3	Distances and affinities between measures 59
§4	The classical concept of absolute continuity 65
*§5	Vitali covering lemma 68
*§6	Densities as almost sure derivatives 70
§7	Problems 71
§8	Notes 75
CHAPTER 4: PRODUCT SPACES AND INDEPENDENCE	
§1	Independence 77
§2	Independence of sigma-fields 80
§3	Construction of measures on a product space 83
§4	Product measures 88
*§5	Beyond sigma-finiteness 93
§6	SLLN via blocking 95
*§7	SLLN for identically distributed summands 97
*§8	Infinite product spaces 99

§9	Problems	102
§10	Notes	108
CHAPTER 5: CONDITIONING		
§1	Conditional distributions: the elementary case	111
§2	Conditional distributions: the general case	113
§3	Integration and disintegration	116
§4	Conditional densities	118
*§5	Invariance	121
§6	Kolmogorov's abstract conditional expectation	123
*§7	Sufficiency	128
§8	Problems	131
§9	Notes	135
CHAPTER 6: MARTINGALE ET AL.		
§1	What are they?	138
§2	Stopping times	142
§3	Convergence of positive supermartingales	147
§4	Convergence of submartingales	151
*§5	Proof of the Krickeberg decomposition	152
*§6	Uniform integrability	153
*§7	Reversed martingales	155
*§8	Symmetry and exchangeability	159
§9	Problems	162
§10	Notes	166
CHAPTER 7: CONVERGENCE IN DISTRIBUTION		
§1	Definition and consequences	169
§2	Lindeberg's method for the central limit theorem	176
§3	Multivariate limit theorems	181
§4	Stochastic order symbols	182
*§5	Weakly convergent subsequences	184
§6	Problems	186
§7	Notes	190
CHAPTER 8: FOURIER TRANSFORMS		
§1	Definitions and basic properties	193
§2	Inversion formula	195
§3	A mystery?	198
§4	Convergence in distribution	198
*§5	A martingale central limit theorem	200
§6	Multivariate Fourier transforms	202
*§7	Cramér-Wold without Fourier transforms	203
*§8	The Lévy-Cramér theorem	205
§9	Problems	206
§10	Notes	208

## Contents

ix

## CHAPTER 9: BROWNIAN MOTION

- §1 Prerequisites 211
- §2 Brownian motion and Wiener measure 213
- §3 Existence of Brownian motion 215
- \*§4 Finer properties of sample paths 217
- §5 Strong Markov property 219
- \*§6 Martingale characterizations of Brownian motion 222
- \*§7 Functionals of Brownian motion 226
- \*§8 Option pricing 228
- §9 Problems 230
- §10 Notes 234

## CHAPTER 10: REPRESENTATIONS AND COUPLINGS

- §1 What is coupling? 237
- §2 Almost sure representations 239
- \*§3 Strassen's Theorem 242
- \*§4 The Yurinskii coupling 244
- §5 Quantile coupling of Binomial with normal 248
- §6 Haar coupling—the Hungarian construction 249
- §7 The Komlós-Major-Tusnády coupling 252
- §8 Problems 256
- §9 Notes 258

## CHAPTER 11: EXPONENTIAL TAILS AND THE LAW OF THE ITERATED LOGARITHM

- §1 LIL for normal summands 261
- §2 LIL for bounded summands 264
- \*§3 Kolmogorov's exponential lower bound 266
- \*§4 Identically distributed summands 268
- §5 Problems 271
- §6 Notes 272

## CHAPTER 12: MULTIVARIATE NORMAL DISTRIBUTIONS

- §1 Introduction 274
- \*§2 Fernique's inequality 275
- \*§3 Proof of Fernique's inequality 276
- §4 Gaussian isoperimetric inequality 278
- \*§5 Proof of the isoperimetric inequality 280
- §6 Problems 285
- §7 Notes 287

## APPENDIX A: MEASURES AND INTEGRALS

- §1 Measures and inner measure 289
- §2 Tightness 291
- §3 Countable additivity 292
- §4 Extension to the  $\cap$ -closure 294
- §5 Lebesgue measure 295
- §6 Integral representations 296
- §7 Problems 300
- §8 Notes 300

## APPENDIX B: HILBERT SPACES

§1	Definitions	301
§2	Orthogonal projections	302
§3	Orthonormal bases	303
§4	Series expansions of random processes	305
§5	Problems	306
§6	Notes	306

## APPENDIX C: CONVEXITY

§1	Convex sets and functions	307
§2	One-sided derivatives	308
§3	Integral representations	310
§4	Relative interior of a convex set	312
§5	Separation of convex sets by linear functionals	313
§6	Problems	315
§7	Notes	316

## APPENDIX D: BINOMIAL AND NORMAL DISTRIBUTIONS

§1	Tails of the normal distributions	317
§2	Quantile coupling of Binomial with normal	320
§3	Proof of the approximation theorem	324
§4	Notes	328

## APPENDIX E: MARTINGALES IN CONTINUOUS TIME

§1	Filtrations, sample paths, and stopping times	329
§2	Preservation of martingale properties at stopping times	332
§3	Supermartingales from their rational skeletons	334
§4	The Brownian filtration	336
§5	Problems	338
§6	Notes	338

## APPENDIX F: DISINTEGRATION OF MEASURES

§1	Representation of measures on product spaces	339
§2	Disintegrations with respect to a measurable map	342
§3	Problems	343
§4	Notes	345

INDEX	347
-------	-----



## Preface

This book began life as a set of handwritten notes, distributed to students in my one-semester graduate course on probability theory, a course that had humble aims: to help the students understand results such as the strong law of large numbers, the central limit theorem, conditioning, and some martingale theory. Along the way they could expect to learn a little measure theory and maybe even a smattering of functional analysis, but not as much as they would learn from a course on Measure Theory or Functional Analysis.

In recent years the audience has consisted mainly of graduate students in statistics and economics, most of whom have not studied measure theory. Most of them have no intention of studying measure theory systematically, or of becoming professional probabilists, but they do want to learn some rigorous probability theory—in one semester.

Faced with the reality of an audience that might have neither the time nor the inclination to devote itself completely to my favorite subject, I sought to compress the essentials into a course as self-contained as I could make it. I tried to pack into the first few weeks of the semester a crash course in measure theory, with supplementary exercises and a whirlwind exposition (Appendix A) for the enthusiasts. I tried to eliminate duplication of mathematical effort if it served no useful role. After many years of chopping and compressing, the material that I most wanted to cover all fit into a one-semester course, divided into 25 lectures, each lasting from 60 to 75 minutes. My handwritten notes filled fewer than a hundred pages.

I had every intention of making my little stack of notes into a little book. But I couldn't resist expanding a bit here and a bit there, adding useful reference material, spelling out ideas that I had struggled with on first acquaintance, slipping in extra topics that my students have seemed to need when writing dissertations, and pulling in material from other courses I have taught and neat tricks I have learned from my friends. And soon it wasn't so little any more.

Many of the additions ended up in starred Sections, which contain harder material or topics that can be skipped over without loss of continuity.

My treatment includes a few eccentricities that might upset some of my professional colleagues. My most obvious departure from tradition is in the use of linear functional notation for expectations, an approach I first encountered in books by de Finetti. I attempt to explain the virtues of this notation in the first two Chapters. Another slight novelty—at least for anyone already exposed to the Kolmogorov interpretation of conditional expectations—appears in my treatment of conditioning, in Chapter 5. For many years I have worried about the wide gap between the free-wheeling conditioning calculations of an elementary probability course and the formal manipulations demanded by rigor. I claim that a treatment starting from the idea of conditional distributions offers one way of bridging the gap,

at least for many of the statistical applications of conditioning that have troubled me the most.

The twelve Chapters and six Appendixes contain general explanations, remarks, opinions, and blocks of more formal material. Theorems and Lemmas contain the most important mathematical details. Examples contain gentler, or less formal, explanations and illustrations. Supporting theoretical material is presented either in the form of Exercises, with terse solutions, or as Problems (at the ends of the Chapters) that work step-by-step through material that missed the cutoff as Exercises, Lemmas, or Theorems. Some Problems are routine, to give students an opportunity to digest the ideas in the text without great mental effort; some Problems are hard.

#### A possible one-semester course

Here is a list of the material that I usually try to cover in the one-semester graduate course.

**Chapter 1:** Spend one lecture on why measure theory is worth the effort, using a few of the Examples as illustrations. Introduce de Finetti notation, identifying sets with their indicator functions and writing  $\mathbb{P}$  for both probabilities of sets and expectations of random variables. Mention, very briefly, the fair price Section as an alternative to the frequency interpretation.

**Chapter 2:** Cover the unstarred Sections carefully, but omitting many details from the Examples. Postpone Section 7 until Chapter 3. Postpone Section 8 until Chapter 6. Describe briefly the generating class theorem for functions, from Section 11, without proofs.

**Chapter 3:** Cover Section 1, explaining the connection with the elementary notion of a density. Take a short excursion into Hilbert space (explaining the projection theorem as an extension of the result for Euclidean spaces) before presenting the simple version of Radon-Nikodym. Mention briefly the classical concept of absolute continuity, but give no details. Maybe say something about total variation.

**Chapter 4:** Cover Sections 1 and 2, leaving details of some arguments to the students. Give a reminder about generating classes of functions. Describe the construction of  $\mu \otimes \Lambda$ , only for a finite kernel  $\Lambda$ , via the iterated integral. Cover product measures, using some of the Examples from Section 4. Explain the need for the blocking idea from Section 6, using the Maximal Inequality to preview the idea of a stopping time. Mention the truncation idea behind the version of the SLLN for independent, identically distributed random variables with finite first moments, but skip most of the proof.

**Chapter 5:** Discuss Section 1 carefully. Cover the high points of Sections 2 through 4. (They could be skipped without too much loss of continuity, but I prefer not to move straight into Kolmogorov conditioning.) Cover Section 6.

*Preface*

xiii

**Chapter 6:** Cover Sections 1 through 4, but skipping over some Examples. Characterize uniformly integrable martingales, using Section 6 and some of the material postponed from Section 8 of Chapter 2, unless short of time.

**Chapter 7:** Cover the first four Sections, skipping some of the examples of central limit theorems near the end of Section 2. Downplay multivariate results.

**Chapter 8:** Cover Sections 1, 2, 4, and 6.

If time is left over, cover a topic from the remaining Chapters.

**Acknowledgments**

I am particularly grateful to Richard Gill, who is a model of the constructive critic. His comments repeatedly exposed weaknesses and errors in the manuscript. My colleagues Joe Chang and Marten Wegkamp asked helpful questions while using earlier drafts to teach graduate probability courses. Andries Lenstra provided some important historical references.

Many cohorts of students worked through the notes, revealing points of obscurity and confusion. In particular, Jeankyung Kim, Gheorghe Doros, Daniela Cojocaru, and Peter Radchenko read carefully through several chapters and worked through numerous Problems. Their comments led to a lot of rewriting.

Finally, I thank Lauren Cowles for years of good advice, and for her inexhaustible patience with an author who could never stop tinkering.

David Pollard  
New Haven

*February 2001*