Introduction to Probability

This classroom-tested textbook is an introduction to probability theory, with the right balance between mathematical precision, probabilistic intuition, and concrete applications. *Introduction to Probability* covers the material precisely, while avoiding excessive technical details. After introducing the basic vocabulary of randomness, including events, probabilities, and random variables, the text offers the reader a first glimpse of the major theorems of the subject: the law of large numbers and the central limit theorem. The important probability distributions are introduced organically as they arise from applications. The discrete and continuous sides of probability are treated together to emphasize their similarities. Intended for students with a calculus background, the text teaches not only the nuts and bolts of probability theory and how to solve specific problems, but also why the methods of solution work.

David F. Anderson is a Professor of Mathematics at the University of Wisconsin-Madison. His research focuses on probability theory and stochastic processes, with applications in the biosciences. He is the author of over thirty research articles and a graduate textbook on the stochastic models utilized in cellular biology. He was awarded the inaugural Institute for Mathematics and its Applications (IMA) Prize in Mathematics in 2014, and was named a Vilas Associate by the University of Wisconsin-Madison in 2016.

Timo Seppäläinen is the John and Abigail Van Vleck Chair of Mathematics at the University of Wisconsin-Madison. He is the author of over seventy research papers in probability theory and a graduate textbook on large deviation theory. He is an elected Fellow of the Institute of Mathematical Statistics. He was an IMS Medallion Lecturer in 2014, an invited speaker at the 2014 International Congress of Mathematicians, and a 2015–16 Simons Fellow.

Benedek Valkó is a Professor of Mathematics at the University of Wisconsin-Madison. His research focuses on probability theory, in particular in the study of random matrices and interacting stochastic systems. He has published over thirty research papers. He has won a National Science Foundation (NSF) CAREER award and he was a 2017–18 Simons Fellow.

CAMBRIDGE MATHEMATICAL TEXTBOOKS

Cambridge Mathematical Textbooks is a program of undergraduate and beginning graduate level textbooks for core courses, new courses, and interdisciplinary courses in pure and applied mathematics. These texts provide motivation with plenty of exercises of varying difficulty, interesting examples, modern applications, and unique approaches to the material.

A complete list of books in the series can be found at
www.cambridge.org/mathematics
Recent titles include the following:

Chance, Strategy, and Choice: An Introduction to the Mathematics of Games and Elections, S. B. Smith
Set Theory: A First Course, D. W. Cunningham
Chaotic Dynamics: Fractals, Tilings, and Substitutions, G. R. Goodson
Introduction to Experimental Mathematics, S. Eilers & R. Johansen
A Second Course in Linear Algebra, S. R. Garcia & R. A. Horn
Exploring Mathematics: An Engaging Introduction to Proof, J. Meier & D. Smith
A First Course in Analysis, J. B. Conway
Introduction to Probability, D. F. Anderson, T. Seppäläinen & B. Valkó

Introduction to Probability

DAVID F. ANDERSON
University of Wisconsin-Madison

TIMO SEPPÄLÄINEN
University of Wisconsin-Madison

BENEDEK VALKÓ
University of Wisconsin-Madison

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

4843/24, 2nd Floor, Ansari Road, Daryaganj, Delhi – 110002, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108415859
DOI: 10.1017/9781108235310

First published 2018

Printed in United States of America by Sheridan Books, Inc.

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

Library of Congress Cataloging-in-Publication Data
Names: Anderson, David F., 1978– | Seppäläinen, Timo O., 1961– |
Valkó, Benedek, 1976–.
Title: Introduction to probability / David F. Anderson, University of
Wisconsin, Madison, Timo Seppäläinen, University of
Wisconsin, Madison, Benedek Valkó, University of Wisconsin, Madison.
Description: Cambridge: Cambridge University Press, [2018] | Series:
Cambridge mathematical textbooks | Includes bibliographical
references and index.
Identifiers: LCCN 2017018747 | ISBN 9781108415859
Subjects: LCSH: Probabilities–Textbooks.
Classification: LCC QA273 .A5534 2018 | DDC 519.2–dc23
LC record available at https://lccn.loc.gov/2017018747

ISBN 978-1-108-41585-9 Hardback

To our families

Cambridge University Press
978-1-108-41585-9 — Introduction to Probability
David F. Anderson , Timo Seppäläinen , Benedek Valkó
Frontmatter
More Information

Contents

Preface *page* xi
To the instructor xiv
From gambling to an essential ingredient of modern science and society xvi

Chapter 1 Experiments with random outcomes 1

1.1 Sample spaces and probabilities 1
1.2 Random sampling 4
1.3 Infinitely many outcomes 11
1.4 Consequences of the rules of probability 14
1.5 Random variables: a first look 21
1.6 Finer points ♣ 26
Exercises 29

Chapter 2 Conditional probability and independence 43

2.1 Conditional probability 43
2.2 Bayes' formula 48
2.3 Independence 51
2.4 Independent trials 58
2.5 Further topics on sampling and independence 62
2.6 Finer points ♣ 71
Exercises 72

Chapter 3 Random variables 89

3.1 Probability distributions of random variables 89
3.2 Cumulative distribution function 95
3.3 Expectation 103
3.4 Variance 113
3.5 Gaussian distribution 119
3.6 Finer points ♣ 123
Exercises 126

Cambridge University Press
978-1-108-41585-9 — Introduction to Probability
David F. Anderson , Timo Seppäläinen , Benedek Valkó
Frontmatter
More Information

Chapter 4 Approximations of the binomial distribution 141

4.1 Normal approximation 142
4.2 Law of large numbers 148
4.3 Applications of the normal approximation 149
4.4 Poisson approximation 155
4.5 Exponential distribution 161
4.6 Poisson process ♦ 165
4.7 Finer points ♣ 169
Exercises 171

Chapter 5 Transforms and transformations 181

5.1 Moment generating function 181
5.2 Distribution of a function of a random variable 188
5.3 Finer points ♣ 196
Exercises 197

Chapter 6 Joint distribution of random variables 205

6.1 Joint distribution of discrete random variables 205
6.2 Jointly continuous random variables 212
6.3 Joint distributions and independence 219
6.4 Further multivariate topics ♦ 227
6.5 Finer points ♣ 235
Exercises 236

Chapter 7 Sums and symmetry 247

7.1 Sums of independent random variables 247
7.2 Exchangeable random variables 255
7.3 Poisson process revisited ♦ 261
Exercises 265

Chapter 8 Expectation and variance in the multivariate setting 271

8.1 Linearity of expectation 271
8.2 Expectation and independence 276
8.3 Sums and moment generating functions 282
8.4 Covariance and correlation 284
8.5 The bivariate normal distribution ♦ 294
8.6 Finer points ♣ 296
Exercises 297

Chapter 9 Tail bounds and limit theorems 309

9.1 Estimating tail probabilities 309
9.2 Law of large numbers 313
9.3 Central limit theorem 315
9.4 Monte Carlo method ♦ 318
9.5 Finer points ♣ 320
Exercises 322

Chapter 10 Conditional distribution 329

10.1 Conditional distribution of a discrete random variable 329
10.2 Conditional distribution for jointly continuous random variables 338
10.3 Conditional expectation 346
10.4 Further conditioning topics ♦ 354
10.5 Finer points ♣ 365
Exercises 366

Appendix A Things to know from calculus 379

Appendix B Set notation and operations 380

Appendix C Counting 385

Appendix D Sums, products and series 399

Appendix E Table of values for $\Phi(x)$ 407

Appendix F Table of common probability distributions 408

Answers to selected exercises 411
Bibliography 424
Index 425

Preface

This text is an introduction to the theory of probability with a calculus background. It is intended for classroom use as well as for independent learners and readers. We think of the level of our book as "intermediate" in the following sense. The mathematics is covered as precisely and faithfully as is reasonable and valuable, while avoiding excessive technical details. Two examples of this are as follows.

- The probability model is anchored securely in a sample space and a probability (measure) on it, but recedes to the background after the foundations have been established.
- Random variables are defined precisely as functions on the sample space. This is important to avoid the feeling that a random variable is a vague notion. Once absorbed, this point is not needed for doing calculations.

Short, illuminating proofs are given for many statements but are not emphasized. The main focus of the book is on applying the mathematics to model simple settings with random outcomes and on calculating probabilities and expectations. Introductory probability is a blend of mathematical abstraction and hands-on computation where the mathematical concepts and examples have concrete real-world meaning.

The principles that have guided us in the organization of the book include the following.

(i) We found that the traditional initial segment of a probability course devoted to counting techniques is not the most auspicious beginning. Hence we start with the probability model itself, and counting comes in conjunction with sampling. A systematic treatment of counting techniques is given in an appendix. The instructor can present this in class or assign it to the students.

(ii) Most events are naturally expressed in terms of random variables. Hence we bring the language of random variables into the discussion as quickly as possible.

(iii) One of our goals was an early introduction of the major results of the subject, namely the central limit theorem and the law of large numbers. These are covered for independent Bernoulli random variables in Chapter 4. Preparation for this influenced the selection of topics of the earlier chapters.

(iv) As a unifying feature, we derive the most basic probability distributions from independent trials, either directly or via a limit. This covers the binomial, geometric, normal, Poisson, and exponential distributions.

Many students reading this text will have already been introduced to parts of the material. They might be tempted to solve some of the problems using computational tricks picked up elsewhere. We warn against doing so. The purpose of this text is not just to teach the nuts and bolts of probability theory and how to solve specific problems, but also to teach you *why* the methods of solution work. Only armed with the knowledge of the "why" can you use the theory provided here as a tool that will be amenable to a myriad of applications and situations.

The sections marked with a diamond ♦ are optional topics that can be included in an introductory probability course as time permits and depending on the interests of the instructor and the audience. They can be omitted without loss of continuity.

At the end of most chapters is a section titled *Finer points* on mathematical issues that are usually beyond the scope of an introductory probability book. In the main text the symbol ♣ marks statements that are elaborated in the *Finer points* section of the chapter. In particular, we do not mention measure-theoretic issues in the main text, but explain some of these in the *Finer points* sections. Other topics in the *Finer points* sections include the lack of uniqueness of a density function, the Berry–Esséen error bounds for normal approximation, the weak versus the strong law of large numbers, and the use of matrices in multivariate normal densities. These sections are intended for the interested reader as starting points for further exploration. They can also be helpful to the instructor who does not possess an advanced probability background.

The symbol ▲ is used to mark the end of numbered examples, the end of remarks, and the end of proofs.

There is an exercise section at the end of each chapter. The exercises begin with a small number of warm-up exercises explicitly organized by sections of the chapter. Their purpose is to offer the reader immediate and basic practice after a section has been covered. The subsequent exercises under the heading *Further exercises* contain problems of varying levels of difficulty, including routine ones, but some of these exercises use material from more than one section. Under the heading *Challenging problems* towards the end of the exercise section we have collected problems that may require some creativity or lengthier calculations. But these exercises are still fully accessible with the tools at the student's disposal.

The concrete mathematical prerequisites for reading this book consist of basic set theory and some calculus, namely, a solid foundation in single variable calculus, including sequences and series, and multivariable integration. Appendix A gives a short list of the particular calculus topics used in the text. Appendix B reviews set theory, and Appendix D reviews some infinite series.

Sets are used from the get-go to set up probability models. Both finite and infinite geometric series are used extensively beginning already in Chapter 1. Single variable integration and differentiation are used from Chapter 3 onwards to work with continuous random variables. Computations with the Poisson distribution from Section 4.4 onwards require facility with the Taylor series of e^x. Multiple integrals arrive in Section 6.2 as we begin to compute probabilities and expectations under jointly continuous distributions.

The authors welcome feedback and will maintain a publicly available list of corrections.

We thank numerous anonymous reviewers whose comments made a real difference to the book, students who went through successive versions of the text, and colleagues who used the text and gave us invaluable feedback. Illustrations were produced with Wolfram Mathematica 11.

The authors gratefully acknowledge support from the National Science Foundation, the Simons Foundation, the Army Research Office, and the Wisconsin Alumni Research Foundation.

Madison, Wisconsin July, 2017

David F. Anderson
Timo Seppäläinen
Benedek Valkó

To the instructor

There is more material in the book than can be comfortably covered in one semester at a pace that is accessible to students with varying backgrounds. Hence there is room for choice by the instructor.

The list below includes all sections not marked with a ♦ or a ♣. It outlines one possible 15-week schedule with 150 minutes of class time per week.

- Week 1. Axioms of probability, sampling, review of counting, infinitely many outcomes, review of the geometric series (Sections 1.1–1.3).
- Week 2. Rules of probability, random variables, conditional probability (Sections 1.4–1.5, 2.1).
- Week 3. Bayes' formula, independence, independent trials (Sections 2.2–2.4).
- Week 4. Independent trials, birthday problem, conditional independence, probability distribution of a random variable (Sections 2.4–2.5, 3.1).
- Week 5. Cumulative distribution function, expectation and variance (Sections 3.2–3.4).
- Week 6. Gaussian distribution, normal approximation and law of large numbers for the binomial distribution (Sections 3.5 and 4.1–4.2).
- Week 7. Applications of normal approximation, Poisson approximation, exponential distribution (Sections 4.3–4.5).
- Week 8. Moment generating function, distribution of a function of a random variable (Sections 5.1–5.2).
- Week 9. Joint distributions (Sections 6.1–6.2).
- Week 10. Joint distributions and independence, sums of independent random variables, exchangeability (Sections 6.3 and 7.1–7.2).
- Week 11. Expectations of sums and products, variance of sums (Sections 8.1–8.2).
- Week 12. Sums and moment generating functions, covariance and correlation (Sections 8.3–8.4).
- Week 13. Markov's and Chebyshev's inequalities, law of large numbers, central limit theorem (Sections 9.1–9.3).
- Week 14. Conditional distributions (Sections 10.1–10.3).
- Week 15. Conditional distributions, review (Sections 10.1–10.3).

The authors invest time in the computations with multivariate distributions in the last four chapters. The reason is twofold: this is where the material becomes more interesting and this is preparation for subsequent courses in probability and stochastic processes. The more challenging examples of Chapter 10 in particular require the students to marshal material from almost the entire course. The exercises under *Challenging problems* have been used for bonus problems and honors credit.

Often the Poisson process is not covered in an introductory probability course, and it is left to a subsequent course on stochastic processes. Hence the Poisson process (Sections 4.6 and 7.3) does not appear in the schedule above. One could make the opposite choice of treating the Poisson process thoroughly, with correspondingly less emphasis, for example, on exchangeability (Section 7.2) or on computing expectations with indicator random variables (Section 8.1). Note that the gamma distribution is introduced in Section 4.6 where it elegantly arises from the Poisson process. If Section 4.6 is skipped then Section 7.1 is a natural place to introduce the gamma distribution.

Other optional items include the transformation of a multivariate density function (Section 6.4), the bivariate normal distribution (Section 8.5), and the Monte Carlo method (Section 9.4).

This book can also accommodate instructors who wish to present the material at either a lighter or a more demanding level than what is outlined in the sample schedule above.

For a lighter course the multivariate topics can be de-emphasized with more attention paid to sets, counting, calculus details, and simple probability models.

For a more demanding course, for example for an audience of mathematics majors, the entire book can be covered with emphasis on proofs and the more challenging multistage examples from the second half of the book. These are the kinds of examples where probabilistic reasoning is beautifully on display. Some topics from the *Finer points* sections could also be included.