Preface

This is the companion second volume to my undergraduate text *Fundamentals of Probability: A First Course*. The purpose of my writing this book is to give graduate students, instructors, and researchers in statistics, mathematics, and computer science a lucidly written unique text at the confluence of probability, advanced stochastic processes, statistics, and key tools for machine learning. Numerous topics in probability and stochastic processes of current importance in statistics and machine learning that are widely scattered in the literature in many different specialized books are all brought together under one fold in this book. This is done with an extensive bibliography for each topic, and numerous worked-out examples and exercises. Probability, with all its models, techniques, and its poignant beauty, is an incredibly powerful tool for anyone who deals with data or randomness. The content and the style of this book reflect that philosophy; I emphasize lucidity, a wide background, and the far-reaching applicability of probability in science.

The book starts with a self-contained and fairly complete review of basic probability, and then traverses its way through the classics, to advanced modern topics and tools, including a substantial amount of statistics itself. Because of its nearly encyclopaedic coverage, it can serve as a graduate text for a year-long probability sequence, or for focused short courses on selected topics, for self-study, and as a nearly unique reference for research in statistics, probability, and computer science. It provides an extensive treatment of most of the standard topics in a graduate probability sequence, and integrates them with the basic theory and many examples of several core statistical topics, as well as with some tools of major importance in machine learning. This is done with unusually detailed bibliographies for the reader who wants to dig deeper into a particular topic, and with a huge repertoire of worked-out examples and exercises. The total number of worked-out examples in this book is 423, and the total number of exercises is 808. An instructor can rotate the exercises between semesters, and use them for setting exams, and a student can use them for additional exam preparation and self-study. I believe that the book is unique in its range, unification, bibliographic detail, and its collection of problems and examples.

Topics in core probability, such as distribution theory, asymptotics, Markov chains, martingales, Poisson processes, random walks, and Brownian motion are covered in the first 14 chapters. In these chapters, a reader will also find basic

coverage of such core statistical topics as confidence intervals, likelihood functions, maximum likelihood estimates, posterior densities, sufficiency, hypothesis testing, variance stabilizing transformations, and extreme value theory, all illustrated with many examples. In Chapters 15, 16, and 17, I treat three major topics of great application potential, empirical processes and VC theory, probability metrics, and large deviations. Chapters 18, 19, and 20 are specifically directed to the statistics and machine-learning community, and cover simulation, Markov chain Monte Carlo, the exponential family, bootstrap, the EM algorithm, and kernels.

The book does not make formal use of measure theory. I do not intend to minimize the role of measure theory in a rigorous study of probability. However, I believe that a large amount of probability can be taught, understood, enjoyed, and applied without needing formal use of measure theory. We do it around the world every day. At the same time, some theorems cannot be proved without at least a mention of some measure theory terminology. Even some definitions require a mention of some measure theory notions. I include some unavoidable mention of measure-theoretic terms and results, such as the strong law of large numbers and its proof, the dominated convergence theorem, monotone convergence, Lebesgue measure, and a few others, but only in the advanced chapters in the book.

Following the table of contents, I have suggested some possible courses with different themes using this book. I have also marked the nonroutine and harder exercises in each chapter with an asterisk. Likewise, some specialized sections with reference value have also been marked with an asterisk. Generally, the exercises and the examples come with a caption, so that the reader will immediately know the content of an exercise or an example. The end of the proof of a theorem has been marked by a □ sign.

My deepest gratitude and appreciation are due to Peter Hall. I am lucky that the style and substance of this book are significantly molded by Peter's influence. Out of habit, I sent him the drafts of nearly every chapter as I was finishing them. It didn't matter where exactly he was, I always received his input and gentle suggestions for improvement. I have found Peter to be a concerned and warm friend, teacher, mentor, and guardian, and for this, I am extremely grateful.

Mouli Banerjee, Rabi Bhattacharya, Burgess Davis, Stewart Ethier, Arthur Frazho, Evarist Giné, T. Krishnan, S. N. Lahiri, Wei-Liem Loh, Hyun-Sook Oh, B. V. Rao, Yosi Rinott, Wen-Chi Tsai, Frederi Viens, and Larry Wasserman graciously went over various parts of this book. I am deeply indebted to each of them. Larry Wasserman, in particular, suggested the chapters on empirical processes, VC theory, concentration inequalities, the exponential family, and Markov chain Monte Carlo. The Springer series editors, Peter Bickel, George Casella, Steve Fienberg, and Ingram Olkin have consistently supported my efforts, and I am so very thankful to them. Springer's incoming executive editor Marc Strauss saw through the final production of this book extremely efficiently, and I have much enjoyed working with him. I appreciated Marc's gentility and his thoroughly professional handling of the transition of the production of this book to his oversight. Valerie Greco did an astonishing job of copyediting the book. The presentation, display, and the grammar of the book are substantially better because of the incredible care

and thoughtfulness that she put into correcting my numerous errors. The staff at SPi Technologies, Chennai, India did an astounding and marvelous job of producing this book. Six anonymous reviewers gave extremely gracious and constructive comments, and their input has helped me in various dimensions to make this a better book. Doug Crabill is the greatest computer systems administrator, and with an infectious pleasantness has bailed me out of my stupidity far too many times. I also want to mention my fond memories and deep-rooted feelings for the Indian Statistical Institute, where I had all of my college education. It was just a wonderful place for research, education, and friendships. Nearly everything that I know is due to my years at the Indian Statistical Institute, and for this I am thankful.

This is the third time that I have written a book in contract with John Kimmel. John is much more than a nearly unique person in the publishing world. To me, John epitomizes sensitivity and professionalism, a singular combination. I have now known John for almost six years, and it is very very difficult not to appreciate and admire him a whole lot for his warmth, style, and passion for the subjects of statistics and probability. Ironically, the day that this book entered production, the news came that John was leaving Springer. I will remember John's contribution to my professional growth with enormous respect and appreciation.

West Lafayette, Indiana   Anirban DasGupta