

Preface

From *Probability, For the Enthusiastic Beginner* (Draft version, March 2016)
David Morin, morin@physics.harvard.edu

This book is written for high school and college students learning about probability for the first time. Most of the book is very practical, with a large number of concrete examples and worked-out problems. However, there are also parts that are a bit theoretical (at least for an introductory book), with many mathematical derivations. All in all, if you are looking for a book that serves as a quick reference, this may not be the one for you. But if you are looking for a book that starts at the beginning and derives everything from scratch in a comprehensive manner, then you've come to the right place. In short, this book will appeal to the reader who has a healthy level of enthusiasm for understanding how and why the standard results of probability come about.

Probability is a very accessible (and extremely fun!) subject, packed with challenging problems that don't require substantial background or serious math. The examples in Chapter 2 are a testament to this. Of course, there are plenty of challenging topics in probability that *do* require a more formal background and some heavy-duty math. This will become evident in Chapters 4 and 5 (and the latter part of Chapter 3). However, technically the only math prerequisite for this book is a comfort with algebra. Calculus isn't relied on, although there are a few problems that do involve calculus. These are marked clearly.

All of the problems posed at the ends of the chapters have solutions included. The difficulty is indicated by stars; most problems have two stars. One star means plug and chug, while three stars mean some serious thinking. Be sure to give a solid effort when solving a problem, and don't look at the solution too soon. If you can't solve a problem right away, that's perfectly fine. Just set it aside and come back to it later. It's better to solve a problem later than to read the solution now. If you do eventually need to look at a solution, cover it up with a piece of paper and read one line at a time, to get a hint to get started. Then set the book aside and work things out for real. That way, you can still (mostly) solve it on your own. You will learn a great deal this way. If you instead head right to the solution and read it straight through, you will learn very little.

For instructors using this book as the assigned textbook for a course, a set of homework exercises is posted at www.people.fas.harvard.edu/~djmorin/book.html. A solutions manual is available to instructors upon request. When sending a request, please point to a syllabus and/or webpage for the course.

The outline of this book is as follows. Chapter 1 covers combinatorics, which is the study of how to count things. Counting is critical in probability, because probabilities often come down to counting the number of ways that something can

happen. In Chapter 2 we dive into actual probability. This chapter includes a large number of examples, ranging from coins to cards to four classic problems presented in Section 2.4. Chapter 3 covers expectation values, including the variance and standard deviation. A section on the “sample variance” is included; this is rather mathematical and can be skipped on a first reading. In Chapter 4 we introduce the concept of a continuous distribution and then discuss a number of the more common probability distributions. In Chapter 5 we see how the binomial and Poisson distributions reduce to a Gaussian (or normal) distribution in certain limits. We also discuss the law of large numbers and the central limit theorem. Chapter 6 is somewhat of a stand-alone chapter, covering correlation and regression. Although these topics are usually found in books on statistics, it makes sense to include them here, because all of the framework has been set. Chapter 7 contains six appendices. Appendix C deals with approximations to $(1 + a)^n$ which are critical in the calculations in Chapter 5, Appendix E lists all of the main results we derive in the book, and Appendix F contains a glossary of notation; you may want to refer to this when starting each chapter.

A few informational odds and ends: This book contains many supplementary remarks that are separated off from the main text; these end with a shamrock, “♣.” The letters N , n , and k generally denote integers, while x and t generally denote continuous quantities. Upper-case letters like X denote a random variable, while lower-case letters like x denote the value that the random variable takes. We refer to the normal distribution by its other name, the “Gaussian” distribution. The numerical plots were generated with Mathematica. I will sometimes use “they” as a gender-neutral singular pronoun, in protest of the present failing of the English language. And I will often use an “’s” to indicate the plural of one-letter items (like 6’s on dice rolls). Lastly, we of course take the frequentist approach to probability in this introductory book.

I would particularly like to thank Carey Witkov for meticulously reading through the entire book and offering many valuable suggestions. Joe Swingle provided many helpful comments and sanity checks throughout the writing process. Other friends and colleagues whose input I am grateful for are Jacob Barandes, Sharon Benedict, Joe Blitzstein, Brian Hall, Theresa Morin Hall, Paul Horowitz, Dave Patterson, Alexia Schulz, and Corri Taylor.

Despite careful editing, there is essentially zero probability that this book is error free (as you can show in Problem 4.16!). If anything looks amiss, please check the webpage www.people.fas.harvard.edu/~djmorin/book.html for a list of typos, updates, additional material, etc. And please let me know if you discover something that isn’t already posted. Suggestions are always welcome.

David Morin
Cambridge, MA