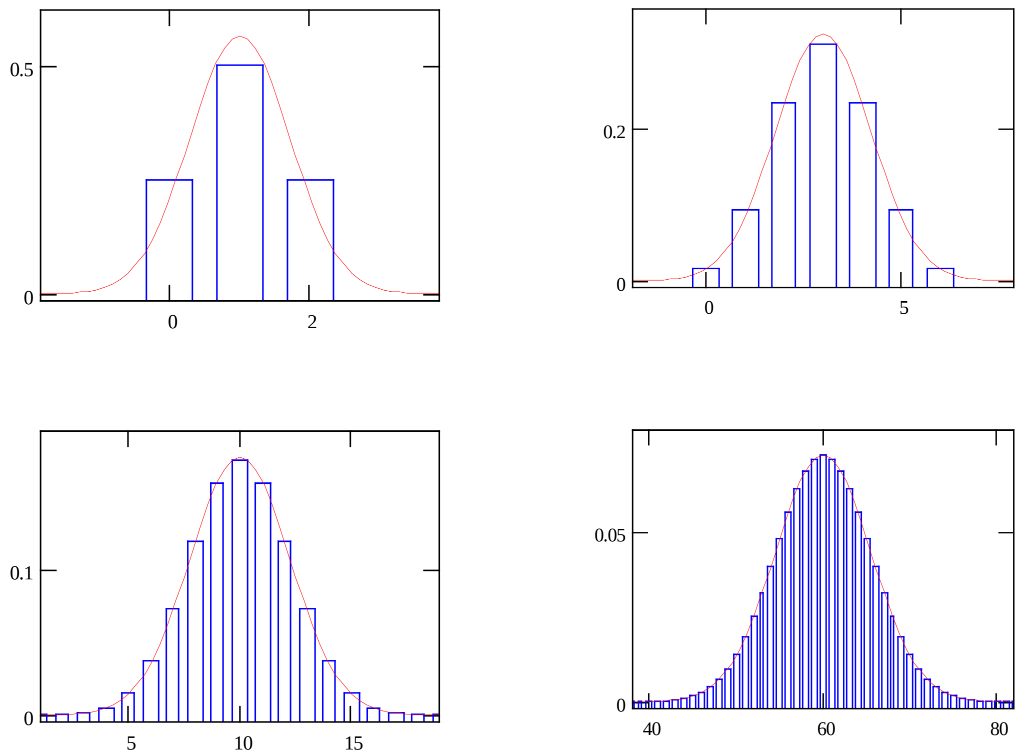
Probability

An Introduction with Applications



Gordon B. Hazen

Preface to the instructor

This text is meant as an introduction to calculus-based probability, and can be used as a preparation for further material in statistics, stochastic processes or decision analysis. It has been used in this fashion in the industrial engineering department at Northwestern University for the past fifteen years. The text has several important features which distinguish it from typical probability texts:

Linguistic presentation of random variables and events

- In this text, I bypass the traditional sample-space/ set-theoretic introduction to probability in favor of an intuitive linguistic presentation involving random variables and events. The set-theoretic foundation for probability is in my view a mathematical nicety which is never practically used in applications and need not appear at the introductory level (although it may be important for more advanced levels of instruction). An intuitive notion of random variable is taken as the primitive, and events are introduced as statements about random variables using the linguistic operators *and*, *or*, *not* rather than the set-theoretic operators of union, intersection, and complement. In my view, this is how experienced users of probability think.

  (For example, when considering whether two events A and B are *disjoint*, an experienced probabilist does *not* think "What is the sample space Ω? What subset of Ω represents A? What subset represents B? Finally, is $A \cap B = \varnothing$?" Rather he/she simply asks "Can A and B occur simultaneously?", which is the way disjointness is presented in this text.) The goal is to move students more quickly towards examples of how probability is *really used*.

- Students are presented early in the course with examples and problems having multiple random variables. This early introduction is easy because, as just stated, random variables are taken as the primitives. Real probability models are *filled* with random variables, both independent and dependent, but typical probability textbooks do not introduce joint distributions until the middle or end of the course. Again, the goal is to move more quickly toward examples of real use of probability.

- Combinatorial probability, that is, the use of counting methods to calculate probabilities, is almost completely bypassed, with the exception of the combination operator $\binom{n}{k}$ for the binomial model. Combinatorial counting problems are peripheral to most important areas of probability and its applications, and the essential features of probability modeling can be presented without including this topic. The usual combinatorial approaches are replaced in this text with approaches based on multiplying conditional probabilities.

Integration of Excel-based Monte Carlo simulation

Spreadsheet-based Monte Carlo simulation is integrated throughout the text, both as a problem solving tool and as an instructional device.

- Early in the text, I introduce a self-contained method for conducting Monte Carlo simulation using spreadsheet software such as Microsoft Excel. No other special software is required, and any reader with spreadsheet software should easily be able to perform a Monte Carlo simulation.
- I use Monte Carlo simulation examples and exercises to help build the reader's intuition about key probability concepts such as the convergence of relative frequencies and long-term averages to probabilities and expected values (the strong law of large numbers). Later on I use Monte Carlo simulation to reinforce intuition about probability density functions for continuous random variables, and to concretely illustrate the implications and meaning of the central limit theorem.
- I emphasize the usefulness of Monte Carlo simulation as a problem-solving tool, especially when algebraic methods are unwieldy. I employ examples and exercises in such applications as electrical power system reliability, hazardous waste transport, inventory modeling, and facility location.
- I discuss the application of elementary techniques of statistical inference to estimation problems arising in Monte Carlo simulation.
- I discuss in depth two important applications in which Monte Carlo simulation is important: *activity networks*, and *probabilistic sensitivity analysis.*

Initial focus on discrete random variables

I present all major concepts first using *discrete* random variables. Later, I present the basics of continuous random variables, noting that all previously introduced properties go through for continuous random variables as long as one replaces pmf by pdf and summation by integral. Again, the goal is to get to applications as quickly as possible without spending time on derivations which are not really used when applying probability tools. I present a table of all major properties in both discrete and continuous forms to assist the student in seeing the parallels.

Emphasis on thinking conditionally

One of the most important skills a student can acquire in preparation for advanced and applied uses of probability is the ability to think *conditionally*. Examples of this manner of thinking include the following.

- Independence is defined in the conditional sense $P(A|B) = P(A)$ rather than the multiplicative sense $P(A \cap B) = P(A)P(B)$. The intuitive view of independence − that *finding out whether* B *occurred does not influence the probability of A* − is emphasized. Too often I have encountered students in advanced courses who know only that independence means you can multiply probabilities. Understanding the conditional definition is crucial to being able to know when to invoke independence assumptions when constructing a probability model.
- Example problems which in many texts are solved by counting (such as calculating the probability of 4-of-a-kind in poker) are solved instead by using sequential multiplication rules for conditional probability.

- A advanced chapter of the text is devoted to conditioning, including the total probability and total expectation rules, conditional independence, conditioning using the expectation operator, the conditional variance formula, and conditional extensions of probability rules.

Emphasis on examples

The text includes many different examples. My method for choosing examples focuses first on finding useful or interesting real situations in which a probability model might be helpful, and only secondly on devising an example which fits the concept currently under discussion. Because real problems are hard, this can often result in examples which are challenging to the novice. My approach is to present solutions in simple concrete (rather than abstract general) ways, and in exercises to have students mimic or incrementally revise solution approaches to examples which would otherwise be too difficult.

Examples and exercises which I present include: Birthday coincidences, airline overbooking, the Windows game Minesweeper, attacking and defending in the board game *Risk*, poker, landslide risk analysis, free-throw shooting, majority voting, baseball batting performance, source-sink network reliability, testing for AIDS, the number of victories by series or round-robin winners, the Illinois lottery, examinations with repeats, evacuating a city, arrivals at an automated entrance gate, single-period inventory models, facility location, electrical power system reliability, and hazardous material transport. I also include a section devoted entirely to activity networks, and another to probabilistic sensitivity analysis.

Use of graphical tools

Throughout the text, I consistently use the *event tree* to give the student visual insight into concepts such as independence, the total probability rule, Bayes' rule, and the binomial and geometric random variables. I also present the material on Poisson splitting graphically using a hybrid transition diagram/ event tree format. I use an intuitive presentation of the *influence diagram* to schematically portray Bayes rule using arrow reversals and to give intuitive meaning to the notion of conditional independence. Overall, I try to be as graphically helpful as possible, presenting bar charts of probability mass functions, plots of cdf's and pdf's, and relative frequency graphs for Monte Carlo simulation whenever possible.

Tools and examples from decision analysis

Event trees and influence diagrams are prominent tools from the field of *decision analysis*. Other decision-analytic tools are also discussed in this text, including sensitivity analysis, tornado diagrams, probabilistic sensitivity analysis, decision trees, expected utility, and of course, Bayes' rule. This text is not intended to be an introduction to decision analysis, but a good preparation for further study can be found here.

Building a student's intuition and highlighting common misconceptions

Throughout the text, but especially in the beginning sections, I make a special effort to build intuition about fundamental notions and warn of common misconceptions. Here are some instances of this practice:

Defining and using random variables

- improperly defining random variables
- confusion of algebraic variables and random variables
- the distribution of a random variable

Conditional probability

- an intuitive view of conditional probability
- conditional probability and temporal order
- conditional probability in a Venn diagram
- misconceptions concerning conditional probability

Independence

- the intuitive view of independence
- the distinction between disjoint events and independent events
- pairwise versus collective independence
- independence is a relationship between random variables
- substituting a conditioning value
- dependent random variables which are uncorrelated
- intuition on independent Bernoulli trials

Expectation and probability operators

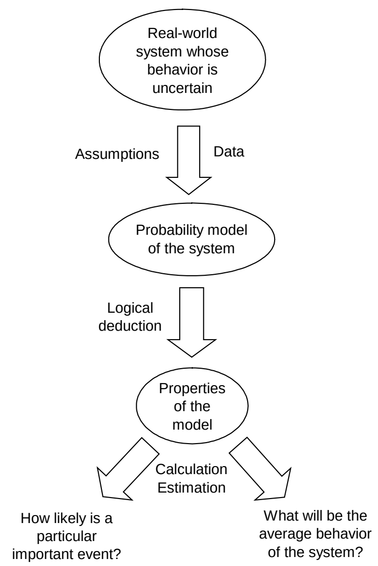
- doing algebra inside the probability operator and expectation operators
- expectation of a function is not the function of the expectation
- mean of a product is not necessarily the product of the means
- the mean of a quotient is not the quotient of the means
- the variance of a sum is not always the sum of the variances

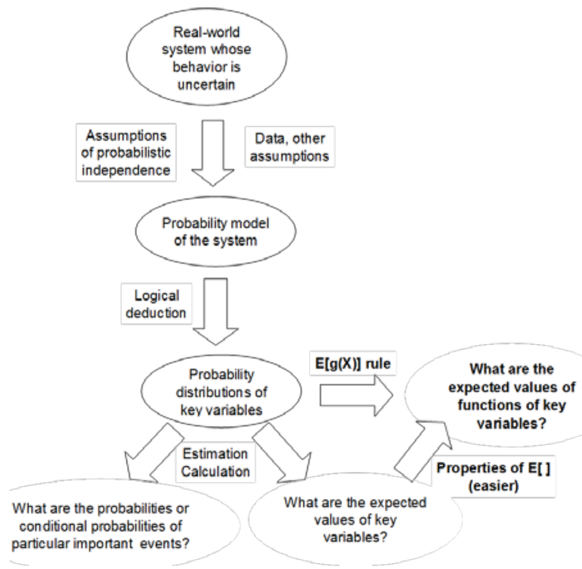
Continuous random variables

- density functions as probability per unit length
- impossible events versus events having probability zero
- relationship between the pdf and the cdf
- lack of memory property for the exponential

Recurring summaries of the role of important topics in probability modeling

The text begins with a preview and overview in the form of a simple flow diagram describing how probability models are used to predict behavior of real-world systems subject to uncertainty. As a review or preview of each basic topic, the probability modeling flow diagram is revisited, and the place of that topic within the diagram is highlighted and described. The purpose is to allow students, after concentrating on the details, to re-focus on the "big picture" of probability modeling.

<table>
<tr>
<th>p. 3, in the overview and preview</th>
<th>p.263, after the student has been introduced to independence, conditional probability, probability distributions, and expectation</th>
</tr>
<tr>
<td>


</td>
<td>


</td>
</tr>
</table>

Applications to stochastic processes, statistics and simulation

A few words about what this text is not. In the concluding chapters, the reader will find short introductions to further topics in which probability modeling plays an important role: Poisson processes, statistical inference, and Monte Carlo simulation. I do not view this text as adequate by itself for courses in these topics. The most important features of Poisson processes are treated, but no further topics in stochastic processes are approached. Only large sample confidence intervals and hypothesis testing are covered, to illustrate the use of the normal distribution and the central limit theorem in statistics. A chapter on Bayesian versus classical statistics is included to give the beginning reader an entry point into this important and timely subject. And only spreadsheet-based Monte Carlo simulation and analysis is treated. For individual courses on any of these topics, there are a variety of appropriate textbooks.

Using this text in a course

I have used this text for a one-quarter (10-week) introductory course in probability in the engeering school at Northwestern University. There is more material here than will fit into a single quarter or semester, and chapters need not be covered in strictly sequential order. Figure A summarizes the precedence relationships between the chapters, and highlights material that is typically included in a one-quarter introduction.

The text contains 360 exercises. Exercise solutions are available on request from the author.

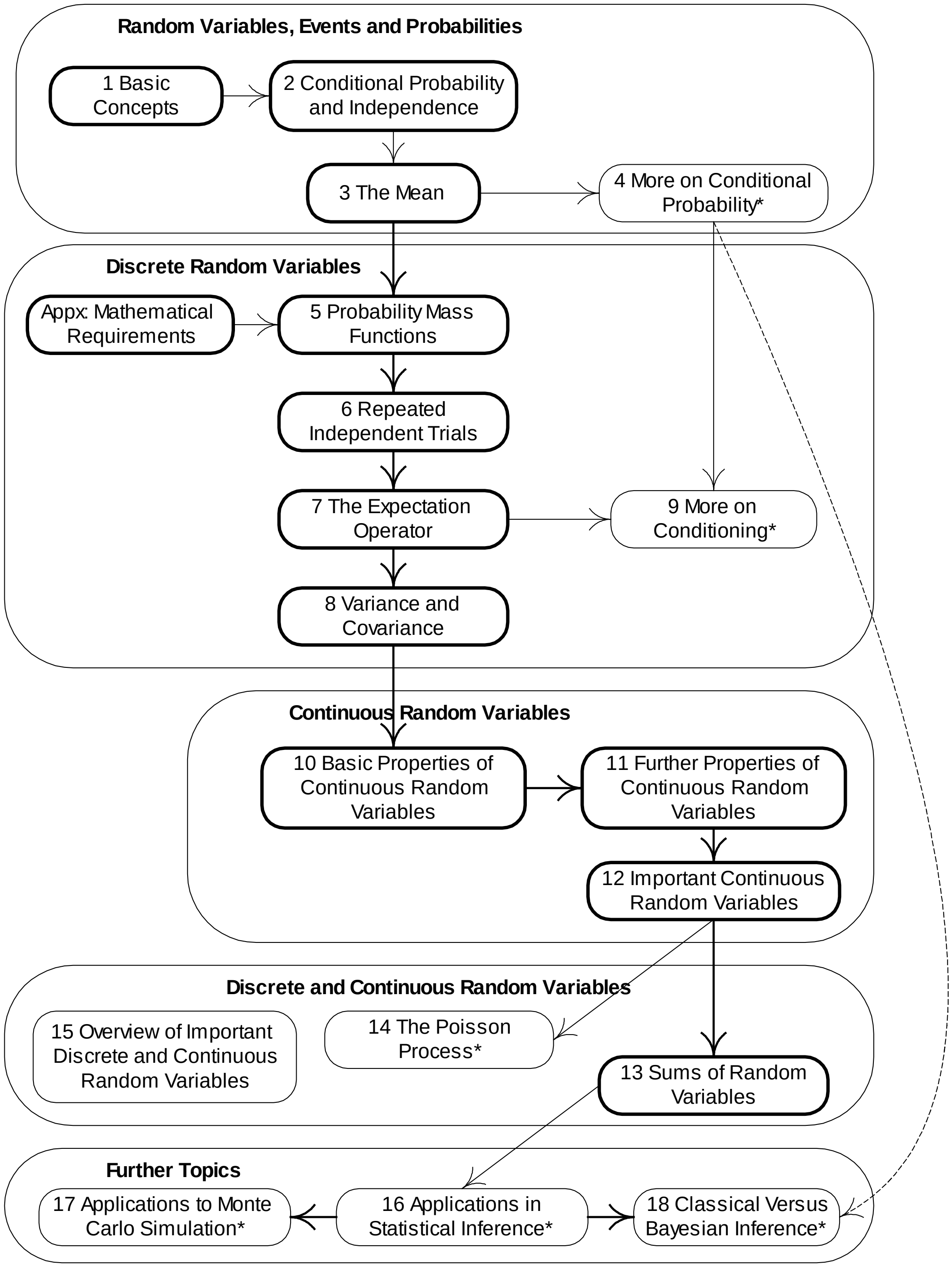


Figure A: Precedence relationships between chapters in this textbook. Arrows indicate what material from prior chapters is used in a given chapter. Dotted arrows indicate that prior material is not heavily used. Chapters outlined in bold form typical topics in a one-quarter or one-semester course. Starred chapters denote optional or advanced topics.

The electronic version of this text

This textbook has been used for fifteen years in the required undergraduate probability course in the Department of Industrial Engineering and Management Sciences at Northwestern University. It is not a supplement, summary, or set of lecture notes – it is the complete version of a 591-page text, viewable on your computer or electronic device. The text is hyperlinked both to and from its table of contents, and also contains crosslinks within its body.

The author

Gordon Hazen is an emeritus professor in the Department of Industrial Engineering and Management Sciences at Northwestern University. He has taught courses in probability, statistics, stochastic processes, and decision analysis for three decades. His research focuses on stochastic and decision analytic applications in healthcare, and he has had extensive editorial service in decision analysis, medical decision making, and operations research.

Contents of the text