PROBABILITY, RANDOM VARIABLES, AND RANDOM PROCESSES

Theory and Signal Processing Applications

JOHN J. SHYNK

WILEY

PROBABILITY, RANDOM VARIABLES, AND RANDOM PROCESSES

Theory and Signal Processing Applications

JOHN J. SHYNK

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA, USA

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

Cover image: © loops7/iStockphoto

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Shynk, John Joseph.
Probability, random variables, and random processes : theory and signal processing applications / John J. Shynk.
p. cm.
Includes index.
ISBN 978-0-470-24209-4 (hardback)
1. Probabilities–Textbooks. 2. Stochastic processes–Textbooks. 3. Engineering–Statistical methods–Textbooks.
I. Title.
QA274.S53 2012
519.2–dc23

2012008186

Printed in the United States of America

ISBN: 9780470242094

10 9 8 7 6 5 4 3 2 1

PROBABILITY, RANDOM VARIABLES, AND RANDOM PROCESSES

To Tokie
In memory of A. V. and E. F.

PREFACE

The goal of this book is to provide a more rigorous mathematical framework for probability, random variables, and random processes than is generally available in most undergraduate textbooks on probability and statistics for engineers. The material is designed for first-year graduate students, though much of it is accessible to seniors who have a strong mathematical background, and if some of the more theoretical sections are excluded by the instructor. The book has several features:

- Numerous detailed figures and tables are given that summarize the various techniques and show example results. These include a combination of more than 600 illustrations and plots generated using MATLAB® that are designed to reinforce the material.
- Many examples are included to illuminate subtle points that are often overlooked or not fully explained in basic textbooks. Each chapter also contains homework problems, for which there is a solution manual available to the instructor.
- Several appendices provide background material on related topics in mathematics and signals and systems so that the book is relatively self-contained. A summary of several parametric univariate distributions is included in Appendix A.
- The third part of the book describes signal processing and communications applications that a student is likely to encounter in subsequent courses in engineering. This introductory material is based on several courses offered at the University of California, Santa Barbara, and serves as a preview of applications that utilize stochastic techniques.

After the introduction in Chapter 1, which includes an overview of the material as well as a review of linear systems and frequency-domain transforms, the book is divided into three parts, each containing four chapters:

- Part I: Probability Theory; Random Variables; Multiple Random Variables; and Expectation and Moments (Chapters 2–5).
- Part II: Random Processes; Stochastic Convergence, Calculus, and Decompositions; Systems, Noise, and Spectrum Estimation; and Sufficient Statistics and Parameter Estimation (Chapters 6–9).
- Part III: Communication Systems and Information Theory; Optimal Filtering; Adaptive Filtering; and Equalization, Beamforming, and Direction Finding (Chapters 10–13). These four chapters are located at www.wiley.com/go/randomprocesses.

Chapter 2 covers basic probability with an emphasis on discrete experiments. Sample spaces, events, and fields are introduced in order to provide a framework for the abstract probability space $\{\Omega, \mathcal{F}, P\}$, which is used later for random variables and random processes. Random variables are defined in Chapter 3, which includes a description of many well-known (and not so well-known) continuous and discrete parametric families of distributions. Multiple random variables are considered in Chapter 4 where several techniques for deriving the distributions of transformations of random variables are investigated. Some multivariate distributions are also included. Chapter 5 defines the expectation of a random variable, as well as expectations of functions of random variables, moments, and the characteristic function. Conditional expectation and several of its important properties are also discussed.

By extending the characterization of multiple random variables to be indexed by time, we introduce random processes in Chapter 6. Various properties of a random process are covered, such as independence and stationarity, and we describe different types of random processes, including independent sequences, Markov chains, and martingales. Specific well-known random processes such as the Poisson and Wiener processes are developed and investigated. Chapter 7 explores other characterizations of random processes, including stochastic continuity, derivatives, integrals, and differential equations. Stochastic convergence of random sequences is also described, as well as the laws of large numbers and the central limit theorem. The power spectral density of a random process is defined in Chapter 8, which is used to describe the behavior of signals when processed (filtered) by a system. We focus on linear time-invariant systems, though some nonlinear processing is mentioned. Spectral estimation using parametric and nonparametric techniques are also discussed. Chapter 9 introduces sufficient statistics and describes several important methods for estimating the parameters of a random variable. The techniques are based on various criteria, including the mean-square error, maximum likelihood, and least squares.

The final part of the book begins with an overview of digital communications in Chapter 10, which includes an introduction to information theory. Detectors based on the maximum a posteriori and maximum likelihood criteria are derived. Optimal filtering techniques are covered in Chapter 11, focusing on the mean-square-error criterion, resulting in causal and noncausal Wiener filters. Linear prediction using a lattice filter, and Kalman filtering based on a state-space signal model are also introduced. Adaptive filtering algorithms and structures are described in Chapter 12, beginning with a discussion of steepest descent and Newton's method. Stochastic methods for studying the convergence of adaptive algorithms and their steady-state properties are described. Finally, adaptive beamforming is introduced in Chapter 13 where multiple antennas are used to collect and separate cochannel signals. Adaptive equalization is also covered, which is used to compensate for signal distortion in a transmission channel. We describe ideal and training-based methods as well as "blind" algorithms that do not require training or a pilot signal. Direction-finding algorithms for estimating the angles of arrival of the signals impinging on the antennas are also introduced.

Seven appendices provide additional background material for the topics presented in this book. They include the following:

- Summaries of univariate distributions for 22 continuous and 11 discrete random variables.
- Continuity of a function and descriptions of several functions with specific symbols used throughout the book.
- Discrete- and continuous-time frequency-domain transforms, with tables of properties and some transform pairs.
- A review of Riemann integration, a brief description of Riemann–Stieltjes and Lebesgue integrals, and a summary of useful indefinite and definite integrals.
- Identities and infinite series, mainly for discrete random variables and random sequences.
- Derivations of inequalities and bounds for expectations, such as Markov's and Chebychev's inequalities and the Cramér–Rao lower bound.
- Several matrix properties including subspaces, decompositions, and differentiation with respect to vectors.

The reader might find it useful to refer to the material in these appendices during the course of instruction, and as a self-contained review of topics from previous courses on calculus, signals and systems, and linear algebra.

The book has been designed so that it could be used for an entire academic year with each quarter covering one of the three parts mentioned earlier:

- *Quarter system.* Fall: Chapters 1–5; Winter: Chapters 6–9; Spring: Chapters 10–13.

For a semester system, it could be divided such that the material on systems, estimation, and the applications are covered during the second semester:

- *Semester system.* Fall: Chapters 1–7; Spring: Chapters 8–13.

However, the book is also easily partitioned for one quarter, two quarters, or one semester by omitting some of the advanced material in several of the chapters. For a one-quarter course (10 weeks of instruction), the instructor should be able to cover much of Chapters 1–8 by omitting, for example, stochastic calculus and spectrum estimation. For a one-semester course (15 weeks of instruction), Chapter 9 on sufficient statistics and parameter estimation might be included, again by omitting some of the earlier material. The application chapters have also been provided for interested students as a preview of material usually covered later in other engineering courses. The instructor may find it appropriate to use some of those topics as examples of systems and the signals they process when presenting the material in earlier chapters on random processes, systems, and noise.

I would like to thank S. Chandrasekaran for reviewing material in the appendices, and J. D. Gibson for providing support from the College of Engineering. I am indebted to my students in ECE 235, ECE 240A, and ECE 245, whose questions have contributed to my appreciation of the subtler aspects of several topics covered in the book. They have provided valuable feedback that led to some of the discussions and illustrative examples. I would also like to thank my colleagues in the Department of Electrical and Computer Engineering, whose interactions and insights over the years have provided me with a deeper understanding of statistical signal processing and the wide range of applications. Finally, thanks to my publisher at Wiley, George Telecki, for encouraging this project, and to Kari Capone, Dan Timek, Stephanie Loh, and Shalini Sharma for their assistance during the final stages of production.

JJS
Santa Barbara, CA
July 2012

NOTATION

Since a wide range of material is covered in this book, we present a brief overview of the notation. In many books on signal processing, the same symbol is used to denote a random process and a realization of that process, which is a deterministic waveform. In this book, we use instead the notation typical of books on probability and random processes:

- An uppercase letter denotes random variable X, random process $X(t)$, or random sequence $X[k]$, where t is continuous time and k is discrete time.
- Lowercase letter x is an outcome of X, $x(t)$ is a realization (continuous waveform) of $X(t)$, and $x[k]$ is a realization (sequence of numbers) of $X[k]$.

Typically, these letters are from the end of the Latin alphabet. One exception to the above notation is that uppercase K, M, and N usually denote (nonrandom) integers, as in the following sums of random variables:

$$\sum_{m=1}^{M} X_m, \qquad \sum_{n=1}^{N} Y_n. \tag{1}$$

If $\{K, M, N\}$ turn out to be random variables in a particular problem such as in a random sum, it will be specifically mentioned.

- A bold uppercase letter denotes random vector $\mathbf{X}$, random vector process $\mathbf{X}(t)$, or random vector sequence $\mathbf{X}[k]$.
- Bold lowercase letter $\mathbf{x}$ is a vector outcome of $\mathbf{X}$, $\mathbf{x}(t)$ is a vector realization (vector of waveforms) of $\mathbf{X}(t)$, and $\mathbf{x}[k]$ is a vector realization (vector of sequences) of $\mathbf{X}[k]$.

All vectors in this book are *column vectors*. A row vector is obtained via transpose $\mathbf{x}^T$ or complex conjugate transpose $\mathbf{x}^H$. The superscript on $\mathbf{x}^*$ denotes only complex conjugation, and does not include transpose.

- A bold uppercase letter $\mathbf{A}$ is also used for a nonrandom matrix, and a bold lowercase letter $\mathbf{a}$ for a nonrandom vector. Typically, these letters are from the beginning of the Latin alphabet.

The reader should be able to determine from the context of a discussion if $\mathbf{X}$ is a random vector or a nonrandom matrix. Two important cases are the autocorrelation matrix $\mathbf{R_{XX}}$ and autocovariance matrix $\mathbf{C_{XX}}$ which are nonrandom quantities for random vector $\mathbf{X}$.

- Calligraphic $\mathcal{E}$ is used for expectation. For example, the autocorrelation matrix above is $\mathbf{R_{XX}} \triangleq \mathcal{E}[\mathbf{X}\mathbf{X}^T]$.

Although E is used for expectation in many books on probability, we use $\mathcal{E}$ because it is necessary in some chapters that E represent an error random variable (E is also used to represent an event in the sample space Ω).

In order to be concise in many equations throughout the book, we have written expressions like $1/2\pi j$, for example, where it is understood that all terms of $2\pi j$ are in the denominator without having to use parentheses $1/(2\pi j)$. Another example is the Gaussian probability density function (pdf):

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(x-\mu)^2/2\sigma^2), \tag{2}$$

where it should be clear that $2\sigma^2$ is in the denominator of the exponent (although that would not be the order of operations in most computer programming languages).

In the glossary at the end of the book, we provide a summary of the notation mentioned here, as well as the following lists of symbols and abbreviations used throughout the book: (i) general symbols and numbers, (ii) Greek symbols (often used for random variable parameters), (iii) calligraphic symbols (for special quantities and expectation $\mathcal{E}$ mentioned above), (iv) mathematical symbols, and (v) abbreviations (acronyms).

CONTENTS