INTRODUCTION TO

PROBABILITY

SECOND EDITION

GEORGE G. ROUSSAS

Introduction to Probability

SECOND EDITION

George G. Roussas

Department of Statistics, University of California, Davis

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK
OXFORD • PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE
SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Table of Contents

Index

Copyright

Academic Press is an imprint of Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
225 Wyman Street, Waltham, MA 02451, USA
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA

Second edition **2014**

Library of Congress Cataloging-in-Publication Data
Roussas, George G., author.
Introduction to probability / by George G. Roussas. -- Second edition.
pages cm
Includes bibliographical references and index.
ISBN 978-0-12-800041-0
1. Probabilities--Textbooks. I. Title.

QA273.R8647 2014
519.2--dc23
2013036799

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

For information on all **Academic Press** publications visit our website at store.elsevier.com

Printed and bound in USA
14 15 16 17 18 10 9 8 7 6 5 4 3 2 1

ISBN: 978-0-12-800041-0

Dedication

To the memory of David Blackwell and Lucien Le Cam

Preface

Overview

This book is an introductory textbook in probability. No prior knowledge in probability is required; however, previous exposure to an elementary precalculus course in probability would prove beneficial in that the student would not see the basic concepts discussed here for the first time.

The mathematical prerequisite is a year of calculus. Familiarity with the basic concepts of linear algebra would also be helpful in certain instances. Often students are exposed to such basic concepts within the calculus framework. Elementary differential and integral calculus will suffice for the majority of the book. In some parts of Chapters 7–11 the concept of a multiple integral is used. In Chapter 11, the student is expected to be at least vaguely familiar with the basic techniques of changing variables in a single or a multiple integral.

Chapter Descriptions

The material discussed in this book is enough for a one-semester course in introductory probability. This would include a rather detailed discussion of Chapters 1–12 except, perhaps, for the derivations of the probability density functions following Definitions 1 and 2 in Chapter 11. It could also include a cursory discussion of Chapter 13.

Most of the material in Chapters 1–12—with a quick description of the basic concepts in Chapter 13—can also be covered in a one-quarter course in introductory probability. In such a case, the instructor would omit the derivations of the probability density functions mentioned above, as well as Sections 9.4, 10.3, 11.3, and 12.2.3.

A chapter-by-chapter description follows. Chapter 1 consists of 16 examples selected from a broad variety of applications. Their purpose is to impress upon the student the breadth of applications of probability, and draw attention to the wide range of situations in which probability questions are pertinent. At this stage, one could not possibly provide answers to the questions posed without the methodology developed in the subsequent chapters. Answers to most of these questions are given in the form of examples and exercises throughout the remaining chapters. In Chapter 2, the concept of a random experiment is introduced, along with related concepts and some fundamental results. The concept of a random variable is also introduced here, along with the basics in counting. Chapter 3 is devoted to the introduction of the concept of probability and the discussion of some basic properties and results, including the distribution of a random variable.

Conditional probability, related results, and independence are covered in Chapter 4. The quantities of expectation, variance, moment-generating function, median, and mode of a random variable are discussed in Chapter 5, along with some basic probability inequalities.

The next chapter, Chapter 6, is devoted to the discussion of some of the commonly used discrete and continuous distributions.

When two random variables are involved, one talks about their joint distribution, as well as marginal and conditional probability density functions and conditional expectation and variance. The relevant material is discussed in Chapter 7. The discussion is pursued in Chapter 8 with the introduction of the concepts of covariance and correlation coefficient of two random variables.

The generalization of concepts in Chapter 7, when more than two random variables are involved, is taken up in Chapter 9, which concludes with the discussion of two popular multivariate distributions and the citation of a third such distribution. Independence of events is suitably carried over to random variables. This is done in Chapter 10, in which some consequences of independence are also discussed. In addition, this chapter includes a result, Theorem 6 in Section 10.3, of significant importance in statistics.

The next chapter, Chapter 11, concerns itself with the problem of determining the distribution of a random variable into which a given random variable is transformed. The same problem is considered when two or more random variables are transformed into a set of new random variables. The relevant results are mostly simply stated, as their justification is based on the change of variables in a single or a multiple integral, which is a calculus problem. The last three sections of the chapter are concerned with three classes of special but important transformations.

The book is essentially concluded with Chapter 12, in which two of the most important results in probability are studied, namely, the weak law of large numbers and the central limit theorem. Some applications of these theorems are presented, and the chapter is concluded with further results that are basically a combination of the weak law of large numbers and the central limit theorem. Not only are these additional results of probabilistic interest, they are also of substantial statistical importance.

As previously mentioned, the last chapter of the book provides an overview of statistical inference.

Features

This book has a number of features that may be summarized as follows. It starts with a brief chapter consisting exclusively of examples that are meant to provide motivation for studying probability.

It lays out a substantial amount of material—organized in twelve chapters—in a logical and consistent manner.

Before entering into the discussion of the concept of probability, it gathers together all needed fundamental concepts and results, including the basics in counting.

The concept of a random variable and its distribution, along with the usual numerical characteristics attached to it, are all introduced early on so that fragmentation in definitions is avoided. Thus, when discussing some special discrete and continuous random variables in Chapter 6, we are also in a position to present their usual numerical characteristics, such as expectation, variance, moment-generating function, etc.

Generalizations of certain concepts from one to more than one random variable and various extensions are split into two parts in order to minimize confusion and enhance understanding. We do these things for two random variables first, then for more than two random variables. Independence of random variables is studied systematically within the framework dictated by the level of the book. In particular, the reproductive property of certain distributions is fully exploited.

All necessary results pertaining to transformation of random variables are gathered together in one chapter, Chapter 11, rather than discussing them in a fragmented manner. This also allows for the justification of the distribution of order statistics as an application of a previously stated theorem. The study of linear transformations provides the tool of establishing Theorem 7 in Section 11.3, a result of great importance in statistical inference.

In Chapter 12, some important limit theorems are discussed, preeminently the weak law of large numbers and the central limit theorem. The strong law of large numbers is not touched upon, as not even an outline of its proof is feasible at the level of an introductory probability textbook.

The book concludes with an overview of the basics in statistical inference. This feature was selected over others, such as elements of Markov chains, of Poisson processes, and so on, in order to provide a window into the popular subject matter of statistics. At any rate, no justice could be done to the discussion of Markov chains, of Poisson processes, and so on, in an introductory textbook.

The book contains more than 150 examples discussed in great detail and more than 450 exercises suitably placed at the end of sections. Also, it contains at least 60 figures and diagrams that facilitate discussion and understanding. In the appendix, one can find a table of selected discrete and continuous distributions, along with some of their numerical characteristics, a table of some formulas used often in the book, a list of some notation and abbreviations, and often extensive answers to the even-numbered exercises.

Concluding Comments

An *Answers Manual*, with extensive discussion of the solutions of all exercises in the book, is available at booksite.elsevier.com/9780128000410 for the instructor.

A table of selected discrete and continuous distributions, along with some of their numerical characteristics, can also be found on the inside covers of the book. Finally, the appendix contains tables for the binomial, Poisson, normal, and chisquare distributions.

The expression log x (logarithm of x), whenever it occurs, always stands for the natural logarithm of x (the logarithm of x with base e). The rule for the use of decimal numbers is that we retain three decimal digits, the last of which is rounded up to the next higher number (if the fourth decimal is greater or equal to 5). An exemption to this rule is made when the division is exact, and when the numbers are read out of tables.

On several occasions, the reader is referred to proofs for more comprehensive treatment of some topics in the book *A Course in Mathematical Statistics*, 2nd edition (1997), Academic Press, by G.G. Roussas.

Thanks are due to my project assistant, Carol Ramirez, for preparing a beautiful set of typed chapters out of a collection of messy manuscripts.

Preface to the Second Edition

This is a revised version of the first edition of the book, copyrighted by Elsevier Inc, 2007.

The revision consists in correcting typographical errors, some factual oversights, rearranging some of the material, and adding a handful of new section exercises.

Also, a substantial number of hints is added to exercises to facilitate their solutions, and extensive cross-references are made as to where several of the examples mentioned are discussed in detail. Furthermore, a brief discussion of the central limit theorem is included early on in Chapter 6, Section 6.2.3. It is done so, because of its great importance in probability theory, and on the suggestion of a reviewer of the book.

Finally, a substantial number of exercises have been added at the end of each chapter. However, because of inadequate time to have them included in the book itself, these exercises will be posted (booksite.elsevier.com/9780128000410), and their answers will be made available to those instructors using this book as a textbook.

In all other respects, the book is unchanged, and the overview, chapter description, features, and guiding comments in the preface of the first edition remain the same.

The revision was skillfully implemented by my associates Michael McAssey and Chu Shing (Randy) Lai, to whom I express herewith my sincere appreciation.

George Roussas

Davis, California

June 2013

CHAPTER 1

Some Motivating Examples

Abstract

Section that is devoted to presenting a number of examples (16 to be precise), drawn from a broad spectrum of human activities. Their purpose is to demonstrate the wide applicability of probability (and statistics). In each case, several relevant questions are posed, which, however, cannot be answered here. Most of these questions are dealt with in subsequent chapters. In the formulation of these examples, certain terms, such as at random, average, data fit by a line, event, probability (estimated probability, probability model), rate of success, sample, and sampling (sample size), are used. These terms are presently to be understood in their everyday sense, and will be defined precisely later on.

Keywords and Phrases

Motivating examples; Landfills; Disease; Circuit; Insurance policy; Extrasensory perception; Unemployment rate; Cancer patients; Radio audience; Production process; Observing the occurrance of an event; Bacteria count; Acidity in rain; Growth of pine seedlings; Comparing medical treatments; Lifetime of an equipment; Human blood types

This chapter consists of a single section that is devoted to presenting a number of examples (16 to be precise), drawn from a broad spectrum of human activities. Their purpose is to demonstrate the wide applicability of probability (and statistics). In each case, several relevant questions are posed, which, however, cannot be answered here. Most of these questions are dealt with in subsequent chapters. In the formulation of these examples, certain terms, such as at random, average, data fit by a line, event, probability (estimated probability, probability model), rate of success, sample, and sampling (sample size), are used. These terms are presently to be understood in their everyday sense, and will be defined precisely later on.

Example 1

In a certain state of the Union, n landfills are classified according to their concentration of three hazardous chemicals: arsenic, barium, and mercury. Suppose that the concentration of each one of the three chemicals is characterized as either high or low. Then some of the questions that can be posed are as follows:

(i) If a landfill is chosen at random from among the n, what is the probability it is of a specific configuration? In particular, what is the probability that it has:

(a) High concentration of barium?

(b) High concentration of mercury and low concentration of both arsenic and barium?

(c) High concentration of any two of the chemicals and low concentration of the third?
(d) High concentration of any one of the chemicals and low concentration of the other two?

(ii) How can one check whether the proportions of the landfills falling into each one of the eight possible configurations (regarding the levels of concentration) agree with a priori stipulated numbers?

(See some brief comments in Chapter 2, and Example 1 in Chapter 3.)

Example 2

Suppose a disease is present in $100p_1\%$ $(0 < p_1 < 1)$ of a population. A diagnostic test is available but is yet to be perfected. The test shows $100p_2\%$ false positives $(0 < p_2 < 1)$ and $100p_3\%$ false negatives $(0 < p_3 < 1)$. That is, for a patient not having the disease, the test shows positive (+) with probability p_2 and negative (−) with probability $1 - p_2$. For a patient having the disease, the test shows "−" with probability p_3 and "+" with probability $1 - p_3$. A person is chosen at random from the target population, is subject to the test, and let D be the event that the person is diseased and N be the event that the person is not diseased. Then, it is clear that some important questions are as follows: In terms of p_1, p_2, and p_3:

(i) Determine the probabilities of the following configurations: D and $+, D$ and $-, N$ and $+, N$ and −.

(ii) Also, determine the probability that a person will test + or the probability the person will test −.

(iii) If the person chosen tests +, what is the probability that he/she is diseased? What is the probability that he/she is diseased, if the person tests −?

(See some brief comments in Chapter 2, and Example 10 in Chapter 4.)

Example 3

In the circuit drawn above (Figure 1.1), suppose that switch $i = 1, \ldots, 5$ turns on with probability p_i and independently of the remaining switches. What is the probability of having current transferred from point *A* to point *B*?

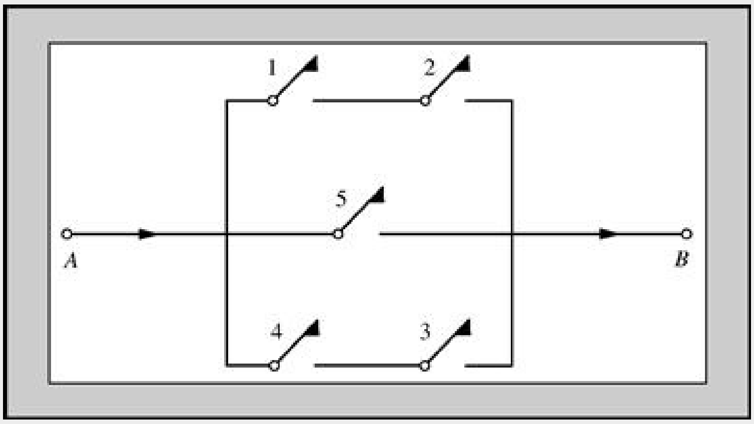

FIGURE 1.1 Transmitting electric current from point A to point B.

(This example is discussed in the framework of Proposition 2(ii) of Chapter 3.)

Example 4

A travel insurance policy pays $1,000 to a customer in case of a loss due to theft or damage on a 5-day trip. If the risk of such a loss is assessed to be 1 in 200, what is a fair premium for this policy?
(This example is discussed in Example 1 of Chapter 5.)

Example 5

Jones claims to have extrasensory perception (ESP). In order to test the claim, a psychologist shows Jones five cards that carry different pictures. Then Jones is blindfolded and the psychologist selects one card and asks Jones to identify the picture. This process is repeated n times. Suppose, in reality, that Jones has no ESP but responds by sheer guesses.

(i) Decide on a suitable probability model describing the number of correct responses.
(ii) What is the probability that at most $n/5$ responses are correct?
(iii) What is the probability that at least $n/2$ responses are correct?

(See a brief discussion in Chapter 2, and Theorem 1 in Chapter 6.)

Example 6

A government agency wishes to assess the prevailing rate of unemployment in a particular county. It is felt that this assessment can be done quickly and effectively by sampling a small fraction n, say, of the labor force in the county. The obvious questions to be considered here are:

(i) What is a suitable probability model describing the number of unemployed?

(ii) What is an estimate of the rate of unemployment?

Example 7

Suppose that, for a particular cancer, chemotherapy provides a 5-year survival rate of 75% if the disease could be detected at an early stage. Suppose further that n patients, diagnosed to have this form of cancer at an early stage, are just starting the chemotherapy. Finally, let X be the number of patients among the n who survive 5 years.

Then the following are some of the relevant questions that can be asked:

(i) What are the possible values of X, and what are the probabilities that each one of these values is taken on?

(ii) What is the probability that X takes values between two specified numbers a and b, say?

(iii) What is the average number of patients to survive 5 years, and what is the variation around this average?

Example 8

An advertisement manager for a radio station claims that over $100p\%$ $(0 < p < 1)$ of all young adults in the city listen to a weekend music program. To establish this conjecture, a random sample of size n is taken from among the target population and those who listen to the weekend music program are counted.

(i) Decide on a suitable probability model describing the number of young adults who listen to the weekend music program.

(ii) On the basis of the collected data, check whether the claim made is supported or not.

(iii) How large a sample size n should be taken to ensure that the estimated average and the true proportion do not differ in absolute value by more than a specified number with prescribed (high) probability?

Example 9

When the output of a production process is stable at an acceptable standard, it is said to be "in control." Suppose that a production process has been in control for some time and that the proportion of defectives has been p. As a means of monitoring the process, the production staff will sample n items. Occurrence of k or more defectives will be considered strong evidence for "out of control."

(i) Decide on a suitable probability model describing the number X of defectives; what are the possible values of X, and what is the probability that each of these values is taken on?

(ii) On the basis of the data collected, check whether or not the process is out of control.

(iii) How large a sample size n should be taken to ensure that the estimated proportion of defectives will not differ in absolute value from the true proportion of defectives by more than a specified quantity with prescribed (high) probability?

Example 10

At a given road intersection, suppose that X is the number of cars passing by until an observer spots a particular make of a car (e.g., a Mercedes).

Then some of the questions one may ask are as follows:

(i) What are the possible values of X?

(ii) What is the probability that each one of these values is taken on?

(iii) How many cars would the observer expect to observe until the first Mercedes appears?

(See a brief discussion in Chapter 2, and Section 6.1.2 in Chapter 6.)

Example 11

A city health department wishes to determine whether the mean bacteria count per unit volume of water at a lake beach is within the safety level of 200. A researcher collected n water samples of unit volume and recorded the bacteria counts.

Relevant questions here are:

(i) What is the appropriate probability model describing the number X of bacteria in a unit volume of water; what are the possible values of X, and what is the probability that each one of these values is taken on?

(ii) Do the data collected indicate that there is no cause for concern?