FOURTH EDITION

Probability and Statistical Inference

ROBERT V. HOGG

ELLIOT A. TANIS

FOURTH EDITION

Probability and Statistical Inference

ROBERT V. HOGG
UNIVERSITY OF IOWA

ELLIOT A. TANIS
HOPE COLLEGE

MACMILLAN PUBLISHING COMPANY
NEW YORK

Maxwell Macmillan Canada
TORONTO

Maxwell Macmillan International
NEW YORK OXFORD SINGAPORE SYDNEY

Editor: Robert W. Pirtle
Production Supervisor: Elaine W. Wetterau
Production Manager: Su Levine
Text and Cover Design: Robert Freese
Illustrations: York Graphic Services, Inc.

This book was set in Times Roman and Optima by York Graphic Services, Inc., printed and bound by Arcata Graphics Company. The cover was printed by Arcata Graphics Company.

Macmillan Publishing Company
866 Third Avenue, New York, New York 10022

Macmillan Publishing Company is part
of the Maxwell Communication Group of Companies.

Maxwell Macmillan Canada, Inc.
1200 Eglinton Avenue East
Suite 200
Don Mills, Ontario M3C 3N1

Library of Congress Cataloging in Publication Data

Hogg, Robert V.
Probability and statistical inference / Robert V. Hogg, Elliot A. Tanis.—4th ed.
p. cm.
Includes index.
ISBN 0-02-355821-0
1. Probabilities. 2. Mathematical statistics. I. Tanis, Elliot A. II. Title.
QA273.H694 1993
519.2—dc20 92-6115
CIP

Printing: 2 3 4 5 6 7 8 Year: 3 4 5 6 7 8 9 0 1

Preface

We are pleased with the reception that was given to the first three editions of *Probability and Statistical Inference*. The fourth edition is still designed for use in a course having from three to six semester hours of credit. No previous study of statistics is assumed, and a standard course in calculus provides an adequate mathematical background. Certain sections have been starred and they are not needed in subsequent sections. This, however, does not mean that these starred sections are unimportant, and we hope many of you will study them.

Although we still view this book as the basis of a junior or senior level course in the mathematics of probability and statistics that many departments of mathematics and/or statistics teach, we have tried to respond to statisticians' recent view of what is good statistical education. In particular, many believe in a data-oriented approach and, if time permits, will assign projects involving the collection of data by teams composed of from four to eight students. Accordingly, Chapter 1 includes good descriptive statistics, exploratory data analysis, the method of least squares, and correlation. Such things as stem-and-leaf displays, box-and-whisker diagrams, quantile-quantile plots, time-sequences (including digidots), and scatter plots are included. These graphical tools are then used throughout the rest of the text.

Chapter 2 provides some basic concepts in probability, much of which is illustrated by the traditional counting problems, although some of this material can be omitted if the instructor so desires. Chapters 3 and 4 contain, respectively, basic discrete and continuous distributions, including their moment-generating functions.

The difficult concept of sampling distribution theory is introduced in Chapter 5 in a very reasonable way. The students learn about the characteristics of the distribution of sums of independent random variables. This, in turn, leads to the Central Limit Theorem and approximations for the binomial distribution, among others. The t- and F-distributions are then introduced. We look upon Section 5.8 as a transition from probability to statistical inference. So often students do not recognize the importance of understanding variability. We point this out and illustrate it with some of the thoughts of W. Edwards Deming and Shewhart's control charts.

Chapter 6 considers confidence intervals and point estimation, including maximum likelihood estimators. In a starred section, the asymptotic distribution of the maximum likelihood estimator is considered. Our discussion of tests of statistical hypotheses in Chapter 7 has been simplified and is more consistent with present-day usage. The considerations of best critical regions and likelihood ratio tests are in starred sections.

Linear models, along with some analysis of variance and regression, are in Chapter 8. Chapter 9 contains multivariate distributions, including the correlation coefficient of two random variables and chi-square goodness of fit tests. Chapter 10 provides a good introduction to nonparametric methods.

In the past many students have asked us to review certain mathematical techniques. We have responded by including Appendix A, which contains helpful notions concerning algebra of sets and certain summations as well as suggestions about limits, series, integration, and multivariate calculus. We believe that we have included some basic mathematical notions here for those who feel that their mathematics is a little "rusty."

Throughout the book, many more figures and real applications have been added. These should help the student understand statistics and what statistical methods can accomplish. For some exercises, it is assumed that calculators or computers are available; thus the solutions will not always involve "nice" numbers. Solutions using a computer are given if a complicated data set is involved. Also, we include answers to almost all of the odd-numbered exercises.

At least 85% of the book could be covered in a two-semester sequence: Probability (much of Chapters 1–5) and Introduction to Mathematical Statistics (much of Chapters 6–10). In a four-semester hour course at the University of Iowa, omitting the starred sections and certain other topics, we cover the first seven chapters plus a few selected topics in the last three chapters. This material could also be covered in a course over two quarters, each with three credit hours.

We are indebted to the *Biometrika* Trustees for permission to include Tables IV and VII, which are abridgments and adaptations of tables published in *Biometrika Tables for Statisticians*. We are also grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to Longman Group Ltd, London, for permission to use Table III from their book *Statistical Tables for Biological, Agricultural, and Medical Research* (6th ed., 1974), reproduced as our Table VI.

We wish to thank our colleagues, students, and friends for many suggestions and for their generosity in supplying data for exercises and examples. Over the years, Elias Ionas has found several corrections; we encourage others to do likewise so that we can correct future printings. Also, we thank Mrs. Julie DeYoung and Mrs. Lori McDowell for their help with the typing, Professor Todd Swanson for his preparation of answers in this book and solutions of even-numbered problems for *The Solutions Manual*, and the University of Iowa and Hope College for providing time, encouragement, and a Faculty Study in Van Wylen Library. Finally, our families, through four editions, have been most understanding during the preparation of all of this material; we truly appreciate their patience and needed their love.

R. V. H.

E. A. T.

Probability and Statistical Inference

Contents

4
Continuous Distributions 191

5
Sampling Distribution Theory 250

6
Estimation 334

7
Tests of Statistical Hypotheses 394

1

Summary and Display of Data

1.1 Basic Concepts

The discipline of statistics deals with the *collection* and *analysis of data*. When measurements are taken, even seemingly under the same conditions, the results usually vary. Despite this variability, a statistician tries to find a pattern; yet due to the "noise," not all of the points lie on the pattern. In the face of this variability, he or she must still do the best to describe the pattern. Accordingly, statisticians know that mistakes will be made in data analysis, and they try to minimize those errors as much as possible and then give bounds on the possible errors. By considering these bounds, decision makers can decide how much confidence they want to place on these data and the analysis of them. If the bounds are wide, perhaps more data should be collected. If they are small, however, the person involved in the study might want to make a decision and proceed accordingly.

Variability is a fact of life, and proper statistical methods can help us understand data collected under inherent variability. Because of this variability, many decisions have to be made that involve uncertainties. In medical research, interest may center on the effectiveness of a new vaccine for mumps; an agronomist must decide if an increase in yield can be attributed to a new strain of wheat; a meteorologist is interested in predicting the probability of rain; the state legislature must decide whether decreasing speed limits will help prevent accidents; the admissions officer of a college must predict the college performance of an incoming freshman; a biologist is interested in estimating the clutch size for a particular type of bird; an economist desires to estimate the unemployment rate; an environmentalist tests whether new controls have resulted in a reduction in pollution.

In reviewing the preceding, relatively short list of possible areas of applications of statistics, the reader should recognize that good statistics is closely associated with careful thinking in many investigations. For illustration, students should appreciate how statistics is used in the endless cycle of the scientific method. We observe Nature and ask questions, we run experiments and collect data that shed light on these questions, we analyze the data and compare the results of the analysis to what we previously thought, we raise new questions, and on and on. Or if you like, statistics is clearly part of the important ''plan–do–study–act'' cycle: Questions are raised and investigations planned and carried out. The resulting data are studied and analyzed and then acted upon, often raising new questions.

There are many aspects of statistics. Some people get interested in the subject by collecting data and trying to make sense out of these observations. In some cases the answers are obvious and little training in statistical methods is necessary. But if a person goes very far in many investigations, he or she soon realizes that there is a need for some theory to help describe the error structure associated with the various estimates of the patterns. That is, at some point appropriate probability and mathematical models are required to make sense out of complicated data sets. Statistics and the probabilistic foundation on which statistical methods are based can provide the models to help people make decisions such as these. So in this book, we are more concerned about the mathematical, rather than the applied, aspects of statistics, although we give enough real examples so that the reader can get a good sense of a number of important applications of statistical methods.

In the study of probability we consider experiments for which the outcome cannot be predicted with certainty. Such experiments are called **random experiments**. Each experiment ends in an outcome that cannot be determined with certainty before the experiment is performed. However, the experiment is such that the collection of every possible outcome can be described and perhaps listed. This collection of all outcomes is called the outcome space or, more frequently, the **sample space S**. The following examples will help illustrate what we mean by random experiments, outcomes, and their associated sample spaces.

Example 1.1-1 A rat is selected at random from a cage, and its sex is determined. The set of possible outcomes is female and male. Thus the sample space is $S = \{\text{female, male}\} = \{F, M\}$.

Example 1.1-2 Each of six students selects an integer at random from the first 52 positive integers. (Or each of the six students selects at random a card from a well-shuffled deck of playing cards.) We are interested in whether at least two of these six integers match (M) or whether they are different (D). Thus $S = \{M, D\}$.

Example 1.1-3 A box of breakfast cereal contains one of four different prizes. The purchase of one box of cereal yields one of the prizes as the outcome, and the sample space is the set of four different prizes.

Example 1.1-4 A state selects a three-digit integer at random for one of its daily lottery games. Each three-digit integer is a possible outcome, and the sample space is $S = \{000, 001, 002, \ldots, 998, 999\}$.

Example 1.1-5 A fair coin is flipped successively at random until the first head is observed. If we let x denote the number of flips of the coin that are required, then $S = \{x : x = 1, 2, 3, 4, \ldots\}$.

Example 1.1-6 A fair coin is flipped successively at random until **heads** is observed on **two** successive flips. If we let y denote the number of flips of the coin that are required, then $S = \{y : y = 2, 3, 4, \ldots\}$.

Example 1.1-7 A biologist is studying certain birds called gallinules that live in a marsh. An adult bird is captured and weighed. If w denotes the weight of the bird, then the sample space would be the set of possible weights. From the biologist's knowledge of gallinules, we could let $S = \{w : 200 \le w \le 450\}$, where w is the weight in grams.

Example 1.1-8 In Example 1.1-7 the biologist could classify a captured bird by sex and weight. In such a case the sample space becomes $S = \{(c, w) : c = \text{F or M}, 200 \le w \le 450\}$, an example of a two-dimensional sample space.

Note that the outcomes of a random experiment can be numerical, as in Examples 1.1-4, 1.1-5, 1.1-6, and 1.1-7, but they do not have to be, as shown by Examples 1.1-1, 1.1-2, and 1.1-3. Often we "mathematize" those latter outcomes by assigning numbers to them; for instance, in Example 1.1-1, we could define a function, say X, such that $X(F) = 0$ and $X(M) = 1$. Such functions, defined on sample spaces, are called **random variables**.

Note the numbers of outcomes in the sample spaces in these examples. In the first four examples each set of possible outcomes is finite. The numbers of outcomes are 2, 2, 4, and 1000, respectively. In Example 1.1-5 the number of possible outcomes is infinite but countable. That is, there are as many outcomes as there are counting numbers (i.e., positive integers). The sample space for Example 1.1-7 is different from the other examples in that the set of possible outcomes is an interval of numbers. Theoretically, the weight could be any one of an infinite number of possible weights; here the number of possible outcomes is not countable. However, from a practical point of view, reported weights are selected from a finite number of possibilities. Many times it is better to conceptualize the sample space as an interval of outcomes and Example 1.1-7 is an example of a sample space of the continuous type.

If we consider a random experiment and its sample space, we note that under repeated performances of the experiment, some outcomes occur more frequently than others. For illustration, in Example 1.1-5, if this coin-flipping experiment is repeated over and over, the first head is observed on the first flip more often than on the second flip. If we can somehow determine the fractions of times a random experiment ends in the respective outcomes, we have described a *distribution* or *population*. Often we cannot determine this distribution through theoretical reasoning but must actually perform the random experiment a number of times to obtain guesses or *estimates* of these fractions. The collection of the observations that are obtained from these repeated trials is often called a *sample*. The making of a conjecture about a distribution based on the sample is called a *statistical inference*. That is, in statistics, we try to argue from the sample to the population. To understand the background behind statistical inferences that are made from the sample, we need a knowledge of some probability, basic distributions, and sampling distribution theory; these topics are considered in the early part of this book. We begin by introducing some terms needed to understand probability.

Given a sample space S, let A be a part of the collection of outcomes in S, that is, $A \subset S$. Then A is called an **event**. When the random experiment is performed and the outcome of the experiment is in A, we say that **event A has occurred**.

We are interested in defining what is meant by the probability of A, denoted by $P(A)$, and often called the chance of A occurring. Sometimes the nature of an experiment is such that the probability of A can be assigned easily. For example, when the state lottery in Example 1.1-4 selects a three-digit integer, we would expect each of the 1000 possible three-digit numbers to have the same chance of being selected, namely 1/1000. If we let $A = \{233, 323, 332\}$, then it makes sense to let $P(A) = 3/1000$. Or if we let $B = \{234, 243, 324, 342, 423, 432\}$, then we would let $P(B) = 6/1000$, the probability of the event B.

Probabilities of events associated with the other examples are perhaps not quite as obvious and straightforward. In Example 1.1-1, the probability of selecting a female rat from the cage depends on the number of female and male rats in the cage. In Example 1.1-2, the probability that at least two students select the same integer is dependent on how randomly the students make their selections.

To help us understand what is meant by the probability of A, $P(A)$, consider repeating the experiment a large number of times, say n times. Count the number of times that event A actually occurred throughout these n performances; this number is called the frequency of event A and is denoted by $\#(A)$. The ratio $\#(A)/n$ is called the **relative frequency** of event A in these n repetitions of the experiment. A relative frequency is usually very unstable for small values of n, but it tends to stabilize as n increases. (You might check this by performing the experiment described in Example 1.1-2, computing the relative frequency of M.) This suggests that we associate with event A a number, say p, that is equal to or approximately equal to the number about which the relative frequency tends to stabilize. This number p can then be taken as the number that the relative frequency of event A will be near in future performances of the experiment. Thus, although we cannot predict the outcome of a random experiment with certainty, we can, for a large value of n, predict fairly accurately the relative frequency associated with event A. The number p assigned to event A is called the **probability** of event A, and it is denoted by $P(A)$. That is, $P(A)$ represents the proportion of outcomes of a random experiment that terminate in the event A in a *large number* of trials of that experiment.

The following two examples will help to illustrate some of the ideas just presented.

Example 1.1-9 Consider the simple experiment of rolling a fair six-sided die (one of a pair of dice) and observing the outcome. The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. If $A = \{1, 2\}$, we would probably let $P(A) = 2/6 = 1/3$. This experiment was simulated 500 times on a computer. We observed the following combinations of the number of trials, n, the frequency of A, $\#(A)$, and the relative frequency of A after n trials, $\#(A)/n$:

n	$\#(A)$	$\#(A)/n$
50	16	0.32
100	34	0.34
250	80	0.32
500	163	0.326

More complete results of the simulation are depicted in Figure 1.1-1. Note that our assignment of 1/3 to $P(A)$ seems to be supported by the simulation.

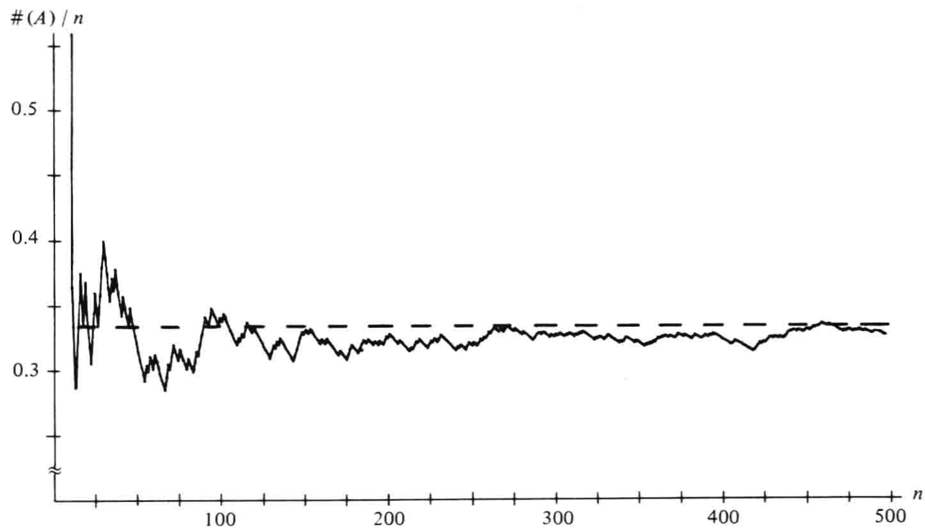

FIGURE 1.1-1

Example 1.1-10 A 12-sided die, called a dodecahedron, has 12 faces that are regular pentagons. These faces are numbered 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12. An experiment consists of rolling this die 12 times. If the face numbered k is the outcome on roll k for $k = 1$ to 12, we say that a match has occurred. The experiment is called a success if at least one match occurs during the 12 trials. Otherwise, the experiment is called a failure. The sample space is $S = \{\text{success, failure}\}$. Let $A = \{\text{success}\}$. We would like to assign a value to $P(A)$. Accordingly, this experiment was also simulated 500 times on a computer. Figure 1.1-2 depicts the results of this simulation, and the following table summarizes a few of those results:

n	$\#(A)$	$\#(A)/n$
50	37	0.74
100	65	0.65
250	166	0.664
500	318	0.636

The probability of event A is not intuitively obvious, but it will be shown in Example 2.4-6 that $P(A) = 1 - (1 - 1/12)^{12} = 0.648$. This assignment is certainly supported by the simulation (although not proved by it).

Examples 1.1-9 and 1.1-10 show that at times intuition can be used to assign probabilities correctly, although at other times probabilities may have to be assigned

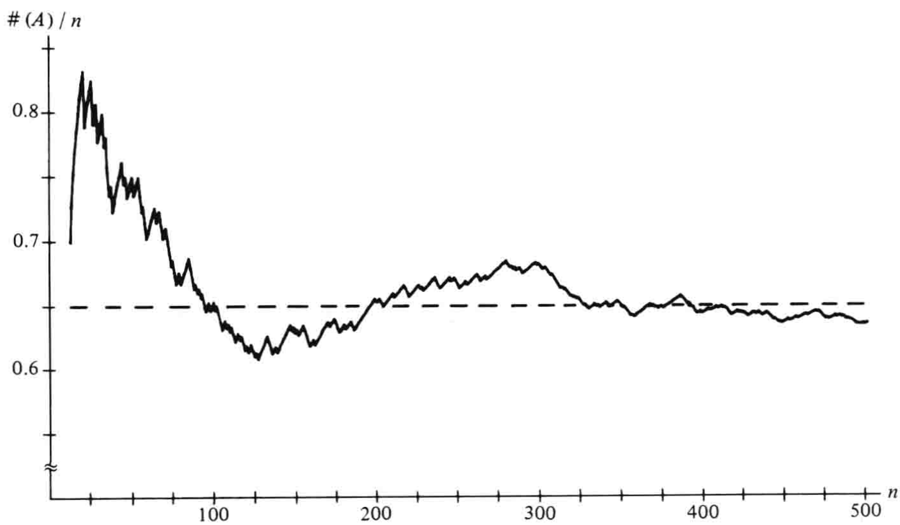

FIGURE 1.1-2

empirically. The following example gives one more illustration where intuition can help in assigning a probability to an event.

Example 1.1-11 A disk 2 inches in diameter is thrown at random on a tiled floor, where each tile is a square with sides 4 inches in length. Let C be the event that the disk will land entirely on one tile. In order to assign a value to $P(C)$, consider the center of the disk. In what region must the center lie to assure that the disk lies entirely on one tile? If you draw a picture, it should be clear that the center must lie within a square having sides of length 2 and lying in the center of a tile. Since the area of this square is 4 and the area of a tile is 16, it makes sense to let $P(C) = 4/16$.

Exercises

1.1-1 Describe the sample space for each of the following experiments:

(a) A student is selected at random from a statistics class, and the student's ACT score in mathematics is determined.

HINT: ACT test scores in mathematics are integers between 1 and 36, inclusive.

(b) A candy bar is selected at random from a production line and is weighed.

(c) A coin is tossed three times, and the sequence of heads and tails is observed.

(d) Some biology students are interested in studying grackles. The sex and wing length of a trapped bird are determined.

1.1-2 Consider families that have three children and select one such family at random. List the outcomes in the following:

(a) The sample space S as 3-tuples, agreeing for example, "gbb" would indicate that the youngest is a girl and the two oldest are boys.

(b) The event $A = \{\text{at least two girls}\}$.

(c) $B = \{\text{exactly one boy}\}$.

(d) $C = \{\text{two youngest are girls}\}$.

(e) Both A and B.

(f) Either B or C (agreeing that this includes being in both B and C).

1.1-3 In each of the following random experiments describe the sample space S. Use your intuition or any experience you may have had to assign a value to the probabilitity p of each of the events A.

(a) The toss of an unbiased coin where the event A is heads.

(b) The cast of an honest die where the event A occurs if we observe a 3, 4, 5, or 6.

(c) The draw of a card from an ordinary deck of playing cards where the event A is a club.

(d) The choice of a point from a square with opposite vertices (0, 0) and (1, 1), where the event A occurs if the sum of the coordinates of the point is less than 3/4.