third edition

Mathematical Statistics

Freund and Walpole

JOHN E. FREUND
Arizona State University

RONALD E. WALPOLE
Roanoke College

mathematical statistics

third edition

Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging in Publication Data

Freund, John E
Mathematical statistics.

Includes bibliographies and index.
1. Mathematical statistics. I. Walpole, Ronald E., joint author. II. Title.
QA276.F692 1980 519.5 79-16146
ISBN 0-13-562066-X

Editorial/production supervision by Karen J. Clemments
Interior design by Judy Winthrop
Cover design by Maurice A. Kruth
Manufacturing buyer: Ed Leone

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA PTY., LIMITED, *Sydney*
PRENTICE-HALL OF CANADA, LTD., *Toronto*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
PRENTICE-HALL OF SOUTHEAST ASIA PTE. LTD., *Singapore*
WHITEHALL BOOKS LIMITED, *Wellington, New Zealand*

preface

Like its first and second editions, this book is designed for a two-semester or three-quarter calculus-based introduction to mathematical statistics. Most of the differences between this edition and the preceding ones reflect the changes that have taken place in recent years in statistical thinking, and in the teaching of statistics. Also, there have been extensive changes in format, which should make the book easier to read and easier to teach.

In addition to substantial changes in notation, the basic material on distribution theory has been reorganized, there is a new chapter combining the material on functions of random variables, the theoretical and applied aspects of estimation have been expanded and placed in two chapters, an expanded coverage is given to nonparametric statistics, the introduction to analysis of variance has been rewritten with more emphasis on the concepts of experimental design, the material on Boolean Algebra has been placed into an appendix, and there are many new exercises and illustrations.

The authors would like to express their appreciation for the many constructive comments which they have received from their colleagues; also, they are indebted to Harry Gaines for his efforts which led to their collaboration on this new edition of MATHEMATICAL STATISTICS, to Ms. Karen J. Clemments for her cooperation during the production stages of the book, to Doug Freund for his editorial assistance, and to Ms. Elizabeth L. Leonard for typing the first draft of the manuscript.

Finally, the authors would like to express their appreciation to the McGraw-Hill Book Company for their permission to reproduce in Table II material from their *Handbook of Probability and Statistics with Tables*, and to Professor E. S. Pearson and the *Biometrika* trustees for their permission to reproduce the material in Tables IV, V, and VI.

John E. Freund

Ronald E. Walpole

contents

PROBABILITY DISTRIBUTIONS

3

MATHEMATICAL EXPECTATION

4

SPECIAL PROBABILITY DISTRIBUTIONS

5

SPECIAL PROBABILITY DENSITIES

6

INTERVAL ESTIMATION

11

HYPOTHESIS TESTING: THEORY

12

HYPOTHESIS TESTING: APPLICATIONS

13

REGRESSION AND CORRELATION

14

ANALYSIS OF VARIANCE

15

NONPARAMETRIC METHODS

16

APPENDIX: THE ALGEBRA OF EVENTS

I

APPENDIX: SUMS AND PRODUCTS

II

introduction

1

1.1 HISTORICAL BACKGROUND

In recent years, the growth of statistics has made itself felt in almost every phase of human activity. Statistics no longer consists merely of the collection of data and their presentation in charts and tables—it is now considered to encompass the science of basing inferences on observed data and the entire problem of making decisions in the face of uncertainty. This covers considerable ground since uncertainties are met when we flip a coin, when a dietician experiments with food additives, when an actuary determines life insurance premiums, when a quality control engineer accepts or rejects manufactured products, when a teacher compares the abilities of his students, when an economist forecasts trends, when a newspaper predicts an election, and so forth.

It would be presumptuous to say that statistics, in its present state of development, can handle all situations involving uncertainties, but new techniques are constantly being developed and modern statistics can, at least, provide the framework for looking at these situations in a logical and systematic fashion. In other words, statistics provides the models that are needed to study situations involving uncertainties, in the same way as calculus provides the models that are needed to describe, say, the concepts of Newtonian physics.

The beginnings of the mathematics of statistics may be found in mid-eighteenth-century studies in probability motivated by interest in games of chance. The theory thus developed for "heads or tails" or "red or black" soon found applications in situations where the outcomes were "boy or girl," "life or death," or "pass or fail," and scholars began to apply probability theory to actuarial problems and some aspects of the social sciences. Later, probability and statistics were introduced into physics by L. Boltzmann, J. Gibbs, and J. Maxwell, and in this century they have found applications in all phases of human endeavor which in some way involve an element of uncertainty or risk. The names which are connected most prominently with the growth of mathematical statistics in the first half of this century are those of R. A. Fisher, J. Neyman, E. S. Pearson, and A. Wald. More recently, the work of R. Schlaifer, L. J. Savage, and others, has given impetus to statistical theories based essentially on methods which date back to the eighteenth-century English clergyman Thomas Bayes.

The approach to statistics presented in this book is essentially the classical approach, with methods of inference based largely on the work of J. Neyman and E. S. Pearson. However, the more general decision-theory approach is introduced in Chapter 9 and some Bayesian methods are presented in Chapter 10.

1.2 MATHEMATICAL PRELIMINARY: COMBINATORIAL METHODS

In many problems of statistics we must list all the alternatives that are possible in a given situation, or at least determine how many different possibilities there are. In connection with the latter, we often use the following theorem, sometimes called the "multiplication rule" for possibilities or choices:

THEOREM 1.1 If an operation consists of two steps, of which the first can be made in n_1 ways and for each of these the second can be made in n_2 ways, then the whole operation can be made in $n_1 \cdot n_2$ ways.

Here, "operation" stands for any kind of procedure, process, or task.

To justify this theorem, let us define the ordered pair (x_i, y_j) to be the outcome which arises when the first step results in possibility x_i and the second step results in possibility y_j. Then, the set of all possible outcomes is composed of the following $n_1 \cdot n_2$ pairs:

$$(x_1, y_1), (x_1, y_2), \ldots, (x_1, y_{n_2})$$

$$(x_2, y_1), (x_2, y_2), \ldots, (x_2, y_{n_2})$$

$$\ldots$$

$$\ldots$$

$$\ldots$$

$$(x_{n_1}, y_1), (x_{n_1}, y_2), \ldots, (x_{n_1}, y_{n_2})$$

EXAMPLE 1.1

Suppose that someone wants to go by bus, by train, or by plane on a week's vacation to one of the five East North Central States. Find the number of different ways in which this can be done.

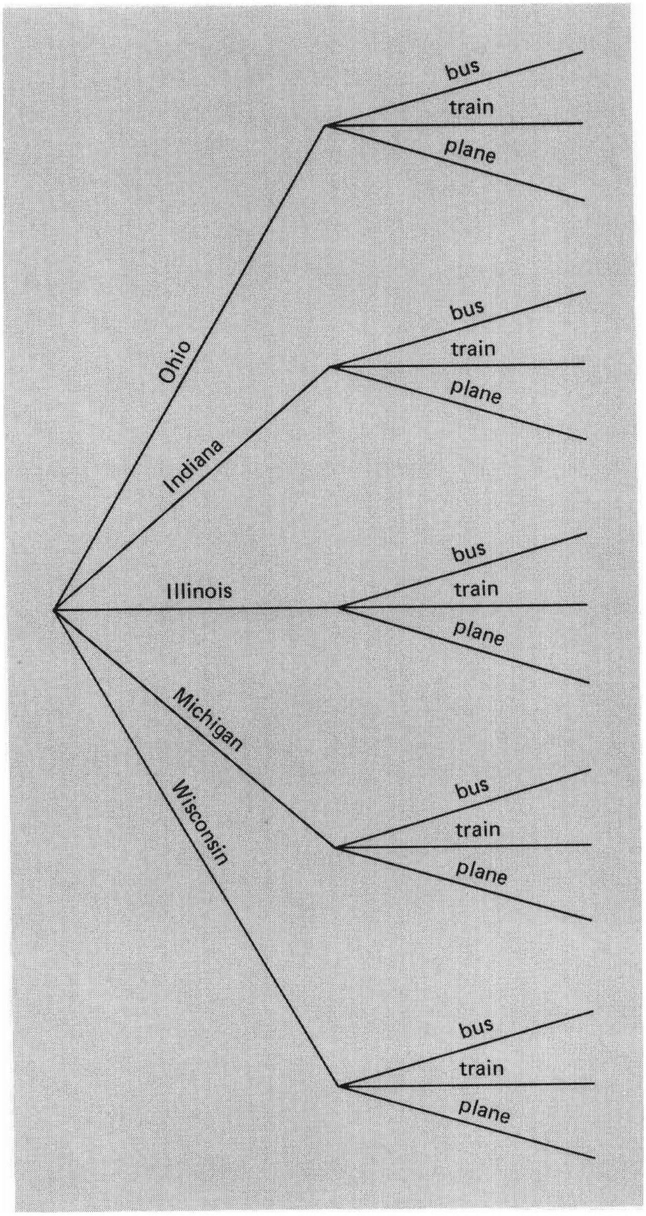

Figure 1.1 Tree diagram.

Solution

The particular state can be chosen in $n_1 = 5$ ways and the means of transportation can be chosen in $n_2 = 3$ ways. Therefore, the trip can be carried out in $5 \cdot 3 = 15$ possible ways. If an actual listing of all the possibilities is desirable, a **tree diagram** like that in Figure 1.1 provides a systematic approach. This diagram shows that there are $n_1 = 5$ branches (possibilities) for the number of states and for each of these branches there are $n_2 = 3$ branches (possibilities) for the different means of transportation. It is apparent that the 15 possible ways of taking the vacation are represented by the 15 distinct paths along the branches of the tree.

EXAMPLE 1.2

How many possible outcomes are there when a red die and a green die are thrown?

Solution

The red die can land in any one of six ways, and for each of these six ways the green die can also land in six ways. Therefore, the pair of dice can land in $6 \cdot 6 = 36$ ways.

Theorem 1.1 may be extended to cover situations where an operation consists of any fixed number of steps. The general case is stated in the following theorem:

THEOREM 1.2 If an operation consists of k steps, of which the first can be made in n_1 ways, for each of these the second step can be made in n_2 ways, for each of the first two the third step can be made in n_3 ways, and so forth, then the whole operation can be made in $n_1 \cdot n_2 \cdot \ldots \cdot n_k$ ways.

EXAMPLE 1.3

How many different lunches are possible consisting of a soup, a sandwich, a dessert, and a drink if one can select from 4 different soups, 3 kinds of sandwiches, 5 desserts, and 4 drinks?

Solution

The total number of lunches would be $4 \cdot 3 \cdot 5 \cdot 4 = 240$.

EXAMPLE 1.4

How many ways can one mark a true-false test consisting of 20 questions?

Solution

If a true-false test consists of 20 questions, there are

$$\underbrace{2 \cdot 2 \cdot 2 \cdot 2 \cdot \ldots \cdot 2 \cdot 2}_{20 \text{ factors}} = 1{,}048{,}576$$

different ways in which one can mark the test, and only one of these corresponds to the case where each answer is correct.

Frequently, we are interested in situations where the outcomes are the different *orders* or *arrangements* that are possible for a group of objects. For example, we might want to know how many different arrangements are possible for electing the president, vice-president, treasurer, and secretary from the 24 members of a club, or we might want to know how many different arrangements are possible for seating 6 persons around a table. Different arrangements like these are called **permutations**.

EXAMPLE 1.5

How many permutations are there of all three of the letters a, b, and c?

Solution

The possible arrangements are *abc*, *acb*, *bac*, *bca*, *cab*, and *cba*, so the number of distinct permutations is six. Using Theorem 1.2, we could have arrived at this answer without actually listing the different permutations. Since there are three choices to select a letter for the first position, then two for the second position, leaving only one letter for the third position, the total number of permutations is $3 \cdot 2 \cdot 1 = 6$.

Generalizing the argument used in this example, we find that n distinct objects can be arranged in $n(n-1)(n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$ ways. We represent this product by the symbol $n!$, which is read "n factorial." Thus, $1! = 1$, $2! = 2 \cdot 1 = 2$, $3! = 3 \cdot 2 \cdot 1 = 6$, and so on. By definition, $0! = 1$.

THEOREM 1.3 The number of permutations of n distinct objects is $n!$.

EXAMPLE 1.6

How many different orders are possible for introducing the 5 starting players of a basketball team to the public?

Solution

There are $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ different orders for introducing the starting lineup.

EXAMPLE 1.7

The number of permutations of the four letters a, b, c, and d is 24, but what is the number of permutations if we take only two of the four letters, or as it is usually put, if we take the four letters two at a time?

Solution

Again using Theorem 1.2, we find that we have two positions to fill with four choices for the first and then three choices for the second for a total of $4 \cdot 3 = 12$ permutations.

Generalizing the argument used in this example, we find that n distinct objects taken r at a time can be arranged in $n(n-1) \cdot \ldots \cdot (n-r+1)$ ways. We represent this product by the symbol ${}_nP_r$.

THEOREM 1.4 The number of permutations of n distinct objects taken r at a time is

$${}_nP_r = n(n-1) \cdot \ldots \cdot (n-r+1)$$

or, in factorial notation,

$${}_nP_r = \frac{n!}{(n-r)!}$$

To obtain the second formula for ${}_nP_r$ we made use of the identity

$$n(n-1) \cdot \ldots \cdot (n-r+1) \cdot (n-r)! = n!$$

In applications, the first formula is generally easier to use, but the one in factorial notation is easier to remember and more easily programmed for solution on a digital computer.

EXAMPLE 1.8

Four names are drawn from the 24 members of a club for the offices of president, vice-president, treasurer, and secretary. In how many different ways can this be done?

Solution

The number of permutations of 24 distinct objects taken 4 at a time is

$$ {}_{24}P_4 = 24 \cdot 23 \cdot 22 \cdot 21 = 255{,}024 $$

EXAMPLE 1.9

In how many ways can a local chapter of the American Chemical Society schedule three speakers for three different meetings, if they are all available on any of five possible dates?

Solution

The number of permutations of 5 distinct objects taken 3 at a time is

$$ {}_5P_3 = 5 \cdot 4 \cdot 3 = 60 $$

Permutations that occur when objects are arranged in a circle are called **circular permutations**. Two circular permutations are not considered different if corresponding objects in the two arrangements are preceded and followed by the same objects as we proceed in a clockwise direction. For example, if four persons are playing bridge, we do not get a new permutation if they all move one position in a clockwise direction.

EXAMPLE 1.10

How many circular permutations are there of four persons playing bridge?

Solution

By considering one person in a fixed position and arranging the other three in 3! ways, we find that there are six different arrangements (circular permutations) of four persons playing bridge.