j. l. hodges, jr. and e. l. lehmann

basic concepts of probability and statistics

second edition

BASIC CONCEPTS OF PROBABILITY AND STATISTICS

J. L. Hodges, Jr. and E. L. Lehmann

University of California, Berkeley

HOLDEN-DAY

San Francisco

Düsseldorf Johannesburg London Mexico Panama

São Paulo Singapore Sydney

TO J. NEYMAN

5 6 7 8 9 0 LC 8 0 7 9 8 7 6 5

500 Sansome Street, San Francisco

Printed in the United States of America.
Library of Congress Catalog Card Number: 72-104973

BASIC CONCEPTS OF PROBABILITY AND STATISTICS

HOLDEN-DAY SERIES IN PROBABILITY AND STATISTICS

E. L. Lehmann, Editor

Bickel & Doksum: Mathematical Statistics
Carlson: Statistics
Freedman: Markov Processes
Freedman: Approximating Countable Markov Chains
Freedman: Brownian Motion and Diffusion
Hájek: Nonparametric Statistics
Hodges & Lehmann: Elements of Finite Probability
Hodges & Lehmann: Basic Concepts of Probability and Statistics
Lehmann: Statistical Methods Based on Ranks: An Introduction to Nonparametric Statistics
Nemenyi: Statistics From Scratch
Neveu: Mathematical Foundations of the Calculus of Probability
Parzen: Stochastic Processes
Rényi: Foundations of Probability
Tanur, Mosteller, Kruskal, Link, Pieters, Rising: Statistics: A Guide to the Unknown

Preface to the Second Edition

The second edition differs from its predecessor primarily through the addition of new material. Stimulated by suggestions of users of the book and by our own teaching experience, we have added more than 300 problems, most of them elementary. We have also provided answers to selected problems at the end of the book. The present volume contains four sections not in the first edition. Of these, Sections 6.9 and 6.10 were introduced in the separately published first part, *Elements of Finite Probability*, in 1965. Section 12.7 discusses the problem of ties for the rank tests treated in Chapter 12, thereby adding to the usefulness of these tests. A final Section 13.6 presents Student's t-test, which many instructors believe should be included in an introductory course, and the Wilcoxon one-sample test. This permits us to compare the t-test with the nonparametric approach emphasized in the other testing material. (Without the calculus the exact distribution of the t-statistic cannot be derived, but we give an approximation for the significance probability.)

We are very grateful to the many readers who have taken the trouble to let us have their reactions and suggestions. These were most helpful as we tried to improve the presentation. In addition to making many minor changes, we rewrote Section 4.1 to correct an error in the earlier version; we should like to thank the several readers who pointed this out to us. We also took a new approach in Section 12.4, in response to a criticism for which we are indebted to Ellen Sherman. Our thanks are due to Howard D'Abrera for checking the problems and providing the answers to the new ones. An answer book can be obtained by instructors from the publishers.

J. L. HODGES, JR.
E. L. LEHMANN

Berkeley
January, 1970

Preface to the First Edition

Statistics has come to play an increasingly important role in such diverse fields as agriculture, criminology, economics, engineering, linguistics, medicine, psychology and sociology. Many of the statistical methods used in these and other areas are quite complicated, and cannot be fully understood without considerable mathematical background. However, the basic concepts underlying the methods require no advanced mathematics. It is the purpose of this book to explain these concepts, assuming only a knowledge of high-school algebra, and to illustrate them on a number of simple but important statistical techniques. We believe that this material is both more interesting and more useful than complicated methods learned by rote and without full understanding of their limitations.

A satisfactory statistical treatment of observational or experimental data requires assumptions about the origin and nature of the data being analyzed. Because of random elements in the data, these assumptions are often of a probabilistic nature. For this reason, statistics rests inherently on probability, and an adequate introduction to statistics, even at the most elementary level, does not seem possible without first developing the essentials of probability theory.

The first part of the book is therefore devoted to the basic concepts of probability. It is centered on the notion of a probability model, which is the mathematical representation of the random aspects of the observations. We attempt not only to develop the essentials of the mathematical theory of probability but also to give a feeling for the relations between the model and the reality which it represents.

Since both probability theory and statistics involve many new concepts which need time for their comprehension, we feel that an introductory course to the subject ideally should extend over more time than the usual one semester or quarter. The 66 sections of the book provide enough material for a year course. However, we have repeatedly used the material in teaching an introduction to probability and statistics in a one-semester course meeting three hours a week. This can be accomplished by taking up only those topics in probability of which essential use is made in the statistics part, omitting Chapter 4 (conditional probability), Chapter 7 (multivariate distributions), as well as certain special topics (Sections 2.4, 3.4, and 6.6–6.8). With this arrangement, the second half of the semester is available for specifically statistical material; here the emphasis may, if desired, be placed either on estimation or on hypothesis testing since the two topics are presented independently of each other.

The conceptual interdependence of different parts of the book, and the resulting possible different course plans, are indicated on the diagram. It should be remarked that certain problems, illustrations and technicalities involve cross-reference outside this schema. However, any of the course plans suggested by the diagram can be used by omitting an occasional illustration or problem, and by accepting a few formulas without proof.

A careful treatment of probability without calculus is possible only if attention is restricted to finite probability models, i.e., to models representing experiments with finitely many outcomes. Fortunately, the basic notions of probability, such as probability model, random variable, expectation and variance, are in fact most easily introduced with finite models. Furthermore, many interesting and important applications lead to finite models, such as the binomial, hypergeometric and sampling models.

In the statistical part of the book, random experimental designs are seen to lead to finite models which, together with the binomial and sampling models and a model for measurement, provide the background for a simple and natural introduction of the basic concepts of estimation and hypothesis testing. Without calculus, it is of course not possible to give a satisfactory treatment of such classical procedures as the one- and two-sample *t*-tests. Instead, we discuss the corresponding Wilcoxon tests, which we believe in any case to be superior.

While basically we have included only concepts that can be defined, and results that can be proved, without calculus, there is one exception: we give approximations to certain probabilities, whose calculation we have explained in principle, but whose actual computation would be too laborious. Thus, in particular, we discuss the normal approximation to the binomial, hypergeometric and Wilcoxon distributions; the Poisson approximation to the binomial and Poisson-binomial distributions, and the chi-square approximation to the goodness-of-fit criterion. Since the limit theorems which underlie these approximations require advanced analytical techniques and are hence not available to us, we give instead numerical illustrations to develop some feeling for their accuracy.

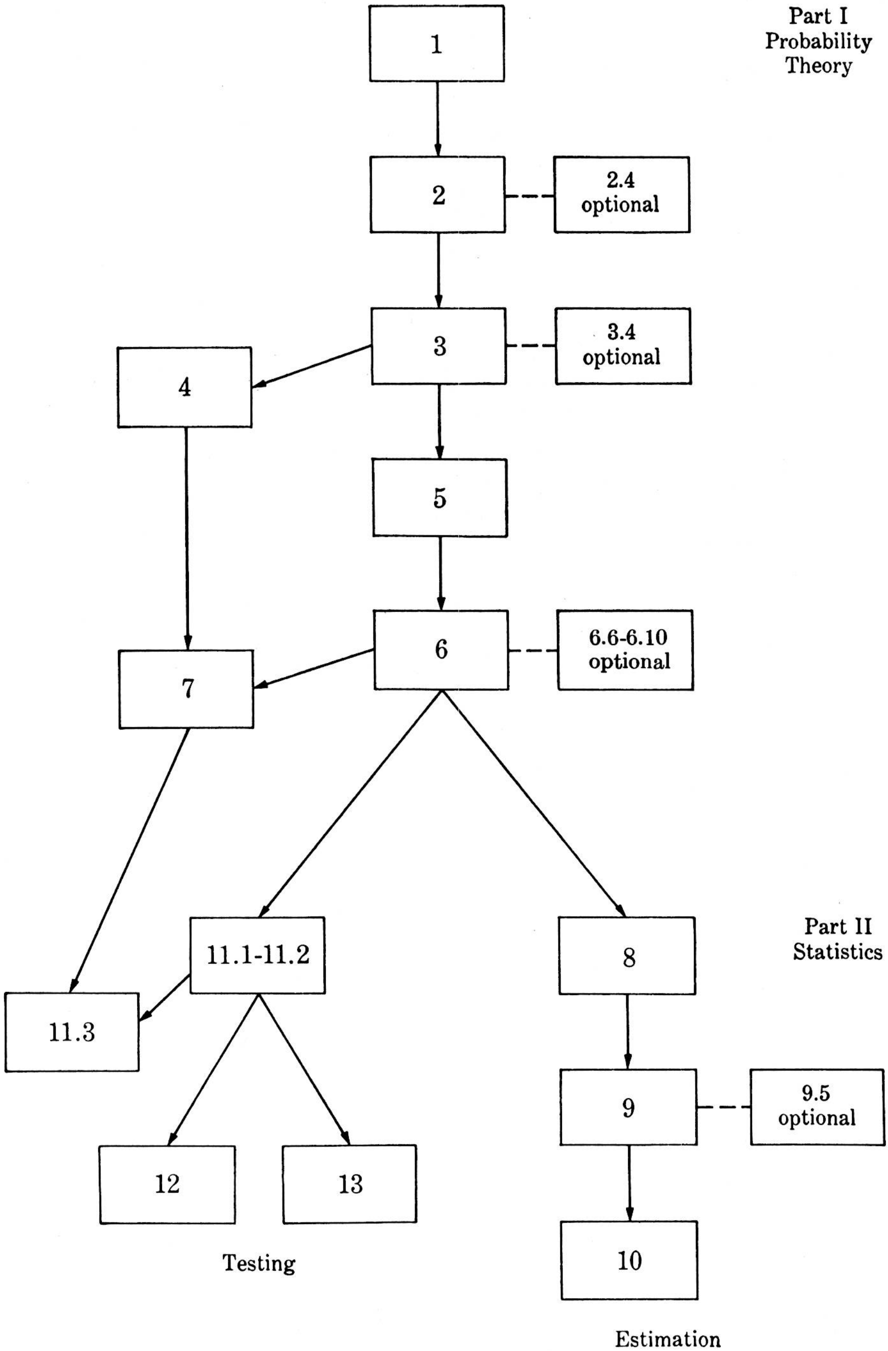
An unusual feature of the book is its careful treatment of the independence assumption in building probability models. Many complex experiments are made up of simpler parts, and it is frequently possible to build models for such experiments by first building simple models for the parts, and then combining them into a "product model" for the experiment as a whole. Particular attention is devoted to the realism of this procedure, and in general to the empirical meaning of the independence assumption, which so often is applied invalidly.

A decimal numbering system is used for sections, formulas, examples, tables, problems, etc. Thus, Section 5.4 is the fourth section of Chapter 5. To simplify the writing, within Chapter 5 itself we omit the chapter and refer to this section simply as Section 4. Similarly Example 5.4.2 is the second example in Section 5.4. However, within Section 5.4 we omit the chapter and section and refer to the example as Example 2; in other sections of Chapter 5 we omit the chapter and refer to the example as Example 4.2. Eleven of the more important tables, including tables for the two Wilcoxon tests, are collected at the end of the book as Tables A–K.

We are grateful to the following colleagues, who have used or read parts of a preliminary edition of the book, for many corrections and suggestions: F. J. Anscombe, M. Atiqullah, D. R. Cox, Tore Dalenius, William Kruskal, L. J. Savage, Henry Scheffé, Rosedith Sitgreaves, Curt Stern, Herbert Solomon and David L. Wallace. Peter Nuesch and Stephen Stigler worked through the problems, and we are indebted to them for numerous corrections. Our thanks are due to Mrs. Julia Rubalcava and Mrs. Carol Rule Roth for patient and expert typing. Finally, we wish to express our appreciation to our publishers, Holden-Day, for their efficiency and helpfulness during all stages of publication.

J. L. Hodges, Jr.
E. L. Lehmann

Berkeley
January, 1964



RELATIONS AMONG THE CHAPTERS

Contents

PART I. PROBABILITY

1. PROBABILITY MODELS

2. SAMPLING

3. PRODUCT MODELS

4. CONDITIONAL PROBABILITY

5. RANDOM VARIABLES

6. SPECIAL DISTRIBUTIONS

7. MULTIVARIATE DISTRIBUTIONS

PART II. STATISTICS

INTRODUCTION TO STATISTICS

8. ESTIMATION

9. ESTIMATION IN MEASUREMENT AND SAMPLING MODELS

10. OPTIMUM METHODS OF ESTIMATION

11. TESTS OF SIGNIFICANCE

12. TESTS FOR COMPARATIVE EXPERIMENTS

13. THE CONCEPT OF POWER

TABLES

PART I
PROBABILITY

CHAPTER 1

PROBABILITY MODELS

1.1 RANDOM EXPERIMENTS

The theories of probability and statistics are mathematical disciplines, which have found important applications in many different fields of human activity. They have extended the scope of scientific method, making it applicable to experiments whose results are not completely determined by the experimental conditions.

The agreement among scientists regarding the validity of most scientific theories rests to a considerable extent on the fact that the experiments on which the theories are based will yield essentially the same results when they are repeated. When a scientist announces a discovery, other scientists in different parts of the world can verify his findings for themselves. Sometimes the results of two workers appear to disagree, but this usually turns out to mean that the experimental conditions were not quite the same in the two cases. If the same results are obtained when an experiment is repeated under the same conditions, we may say that the result is determined by the conditions, or that the experiment is *deterministic*. It is the deterministic nature of science that permits the use of scientific theory for predicting what will be observed under specified conditions.

However, there are also experiments whose results vary, in spite of all efforts to keep the experimental conditions constant. Familiar examples are provided by gambling games: throwing dice, tossing pennies, dealing from a shuffled deck of cards can all be thought of as "experiments" with unpredictable results. More important and interesting instances occur in many fields. For example, seeds that are apparently identical will produce plants of differing height, and repeated weighings of the same object with a chemical balance will show slight variations. A machine which sews together two pieces of material, occasionally—for no apparent reason—will miss a stitch. If we are willing to stretch our idea of "experiment," length of life may be considered a variable experimental result, since people living under similar conditions will die at different and unpredictable ages. We shall refer to experiments that are not deterministic, and thus do not always yield the same result when repeated under the same conditions, as *random experiments.* Probability theory and statistics are the branches of mathematics that have been developed to deal with random experiments.

Let us now consider two random experiments in more detail. As the first example, we take the experiment of throwing a die. This is one of the simplest random experiments and one with which most people are personally acquainted. In fact, probability theory had its beginnings in the study of dice games.

EXAMPLE *1.* *Throwing a die.* Suppose we take a die, shake it vigorously in a dice cup, and throw it against a vertical board so that it bounces onto a table. When the die comes to rest, we observe as the experimental result the number, say X, of points on the upper face. The experiment is not deterministic: the result X may be any of the six numbers 1, 2, 3, 4, 5, or 6, and no one can predict which of the values will be obtained on any particular performance of the experiment. We may make every effort to control or standardize the experimental conditions, by always placing the die in the cup in the same position, always shaking the cup the same number of times, always throwing it against the same spot on the backboard, and so on. In spite of all such efforts, the result will remain variable and unpredictable.

EXAMPLE *2.* *Measuring a distance.* It is desired to determine the distance between two points a few miles apart. A surveying party measures the distance by use of a surveyor's chain. It is found in practice that the measured distance will not be exactly the same if two parties do the job, or even if the same party does the job on consecutive days. In spite of the best efforts to measure precisely, small differences will accumulate and the final measurement will vary from one performance of the experiment to another.

How is it possible for a scientific theory to be based on indeterminacy? The paradox is resolved by an empirical observation: while the result on any particular performance of such an experiment cannot be predicted, a long sequence of performances, taken together, reveals a stability that can serve as the basis for quite precise predictions. The property of *long-run stability* lies at the root of the ideas of probability and statistics, and we shall examine it in more detail in the next section.