Second Edition

Essence of Statistics

Geoffrey R. Loftus
and Elizabeth F. Loftus

SECOND EDITION

ESSENCE OF STATISTICS

Geoffrey R. Loftus
University of Washington

Elizabeth F. Loftus
University of Washington

Alfred A. Knopf Series in Psychology
Series Advisor: Geoffrey Keppel
University of California at Berkeley

ALFRED A. KNOPF  New York

This book is dedicated to Edith L. Loftus and Russell Loftus, without whom the book would have consisted only of Chapter 17.

THIS IS A BORZOI BOOK

PUBLISHED BY ALFRED A. KNOPF, INC.

Second Edition

987654321

Library of Congress Cataloging-in-Publication Data

Loftus, Geoffrey R.
Essence of statistics.

Bibliography: p.
Includes index.
1. Social sciences—Statistical methods.
2. Statistics. I. Loftus, Elizabeth F., 1944–
II. Title.
HA29.L8357 1987 519.5 87-26170
ISBN 0-394-35960-7

Manufactured in the United States of America

Cover Photo: Martin Rogers/Stock, Boston

PREFACE

We wrote the first edition of this book for three reasons. First, we wanted to present the mathematical rationale underlying various statistical concepts in a way that didn't inflict terror on those who were uncomfortable with mathematics. Second, we wanted to range over sufficient territory that various commonly used statistical techniques, such as repeated-measures analysis of variance and planned comparisons, would be included within our scope. Finally, we have found in our courses that a somewhat unusual organization—utilizing an initial heavy emphasis on basic probability and probability distributions—is highly useful in terms of providing an underlying unity for the material that follows. This organization is reflected in both the first and the current edition of the book.

Statistics is fundamentally a mathematical discipline. For the social science student who enjoys working with mathematics, the task of learning statistics can be a source of pleasure and fulfillment. But for the other student—the one frightened by mathematics—the acquisition of statistical knowledge is often accompanied by frustration and pain. We try in this book to cater to both these

students. Our strategy is to present initially a particular statistical concept in its most intuitive form, typically through an appropriate example. It is only after these intuitions are firmly implanted that we unveil the general mathematical machinery underlying them; at that point, the relationships between the intuitions and the mathematics are emphasized and re-emphasized. This strategy constitutes the core of the book.

Finally, however, there are those high-level underpinnings of the material enlightening for the mathematically sophisticated but difficult for the uninitiated, and generally nonessential for basic understanding. These underpinnings are presented as "Digressions" that are collected at the ends of chapters. Disgressions can be read for pleasure and deeper understanding or they can be skipped without a break in the flow of reasoning.

The conceptual design of this book also stems from our observations of the sorts of experimental designs commonly used in the social sciences. Most of these designs, and their associated statistical analyses, are somewhat complex, involving multiple factors and repeated measures, post-hoc tests, planned comparisons, and other associated statistical paraphernalia. One could teach these topics by using a short introductory statistics textbook for the fundamentals and then a higher-level book for the complex material. This strategy, however, requires the student to confront changes in notation with a loss of continuity, and such disruptions cause some degree of disorientation. We prefer the strategy represented in this book where the student travels the entire route using the same text.

We believe that statistics, like any other branch of mathematics, is best viewed and taught as a cumulative discipline, with any given topic building on the foundation of previous topics. That is how this book is organized. We start with the elements of probability theory (Chapter 1) and move on to the general concepts of random variables and probability distributions (Chapter 2). Probability distributions are then used as a means of unifying most of the material in the rest of the book. Most important is a continual emphasis that the process of hypothesis testing doesn't change throughout a variety of tests (sign test, z-test, t-test, F-test, X^2-test). All that changes is the probability distribution of the test statistic under consideration.

This "cumulative philosophy" also dictates a fairly natural order of progression through the various experimental situations and associated statistical tests that we have chosen to include. As we have just suggested, understanding hypothesis testing rests on the understanding of probability distributions, which in turn requires a knowledge of basic probability. Following Chapters 1–3 on these topics, and Chapter 4 on descriptive statistics, we shift to inferential statistics, which constitutes the bulk of the book.

The various experimental situations, along with their associated tests, progress from simple to complex. The first test, described in Chapter 5, is the sign test (based on the binomial distribution), followed in Chapters 6 and 7 by the z-test (based on the normal distribution, which is the continuous version of the binomial). Following two chapters on power and parameter estimation comes

the *t*-test (Chapter 10), which is similar to a *z*-test except that it is used when the variance of the underlying dependent variable is unknown rather than known. Chapter 11, following on the heels of the *t*-test, describes the one-way analysis of variance (ANOVA) situation, portrayed as being similar to a *t*-test situation, but which can incorporate arbitrarily many levels of the independent variable. Along with the ANOVA situation we describe the *F*-test, portrayed as being like a *z*- and a *t*-test in the sense that all three test statistics (*z*, *t*, and *F*) consist of some measure of variance between groups divided by some measure of variance within groups.

After one-way ANOVA comes multifactor ANOVA (Chapter 12), portrayed as being like one-way ANOVA but with more than one independent variable. The major new concept introduced here is that of an *interaction*. In Chapter 13, we introduce repeated-measures designs, which are portrayed as being like multifactor ANOVA but with subjects as one of the factors. Here, we introduce the idea that in a repeated-measures design we are interested in the degree to which subjects are *consistent over the conditions*, and we use this intuition as the basis for describing why the subject-by-condition interaction is used as the error term.

Following Chapter 14 on the rudiments of regression and correlation, we describe in Chapter 15 what to do in addition to (or instead of) an ANOVA. Here we cover post-hoc tests, planned comparisons, and variance accounted for (w^2). In Chapter 16 we describe the X^2 distribution, and various kinds of X^2 tests. Finally, in Chapter 17, we introduce nonparametric tests.

By its very nature, the cumulative philosophy embodied in the book limits flexibility with respect to the order in which topics can be covered. This inflexibility of progression benefits the student who can more easily see how complex concepts are built on a foundation of simpler ones, and thereby ends up viewing statistics as a coherent, unified structure rather than a collection of unrelated subtopics.

But the book is not so structured to make topic order completely immutable. In particular, the topics covered in Chapter 15 (post-hoc tests, planned comparisons, and w^2) could be covered earlier. This chapter was placed in a seemingly inappropriate spot—following a chapter on correlation—so the student can understand that the heart of a planned comparison is an over-condition correlation between a set of weights and a set of means. However, the instructor may wish to cover these topics earlier. Likewise, the topics of power (Chapter 8), parameter estimation (Chapter 9), X^2 (Chapter 16), and nonparametric statistics (Chapter 17) are placed where they seem to fit best, but they can all be taken up earlier in the book should the instructor so desire.

So much for why we organized the book the way we did. Now a few words on how we feel about the way statistical technique is currently applied in social science research. We emphasize and thoroughly describe hypothesis testing. This emphasis reflects the prevalence of hypothesis testing in social science journals and the inescapable requirement that students learn to interpret what they read. However, this emphasis should *not* be taken as an indication that we believe hypothesis testing is the best technique for eliciting interesting information from

a set of data. In fact, we believe that there are a number of analysis techniques that are better for this purpose and, indeed, an increasing stress on these techniques was one of our principal motivations for writing the second edition of this book. These techniques include:

1. A clear initial portrayal of the data through descriptive measures, emphasizing sample statistics, graphical techniques, and associated confidence intervals.
2. A clear assessment of the statistical power of an experiment (best represented by confidence intervals).
3. Clear alternative hypotheses, which are tested by planned comparisons.
4. Delineation of the strength of association between variables, by measures such as r^2 and w^2.

The first technique is stressed throughout the book. Indeed, most experimental situations (e.g., two-way ANOVA) are introduced with an example in which data are shown graphically with associated confidence intervals; based on this portrayal, the major conclusions from the data are tentatively inferred. The ensuing hypothesis-testing procedures are then viewed primarily as confirmatory exercises; advance inspection of the data usually makes their outcomes almost forgone conclusions.

The second edition of this book has benefited tremendously from several people who deserve special thanks. Sarah Lehman orchestrated the seemingly impossible transition from the first edition to the typed manuscript of this version. Leonard Feldman and Walter Nelson provided accurate homework answers; Leonard Feldman also painstakingly inspected the manuscript, tracking, finding, and eliminating errors of all sorts. Geoff Keppel, the editor of this series, did his best to keep us on track when we strayed. Many reviewers—Foster L. Brown, State University of New York, College at Oneonta; Agatha Carroo, North Carolina Central University; Norma Graham, Columbia University; John K. Kennedy, Jr., New York University; George Knight, Arizona State University; and Elke Weber, University of Illinois-Champaign—provided invaluable information about the virtues and drawbacks of the first edition, which we have tried to incorporate here. Brian Wandell, who reviewed the manuscript of the first edition, was his customary inspirational self in providing suggestions for the second. And finally, we express heartfelt appreciation to the people at Random House—Eleanor Castellano, Alison Husting, and especially our editor, Mary Falcon—for putting it all together and putting up with us.

Geoffrey R. Loftus
Elizabeth F. Loftus

CONTENTS