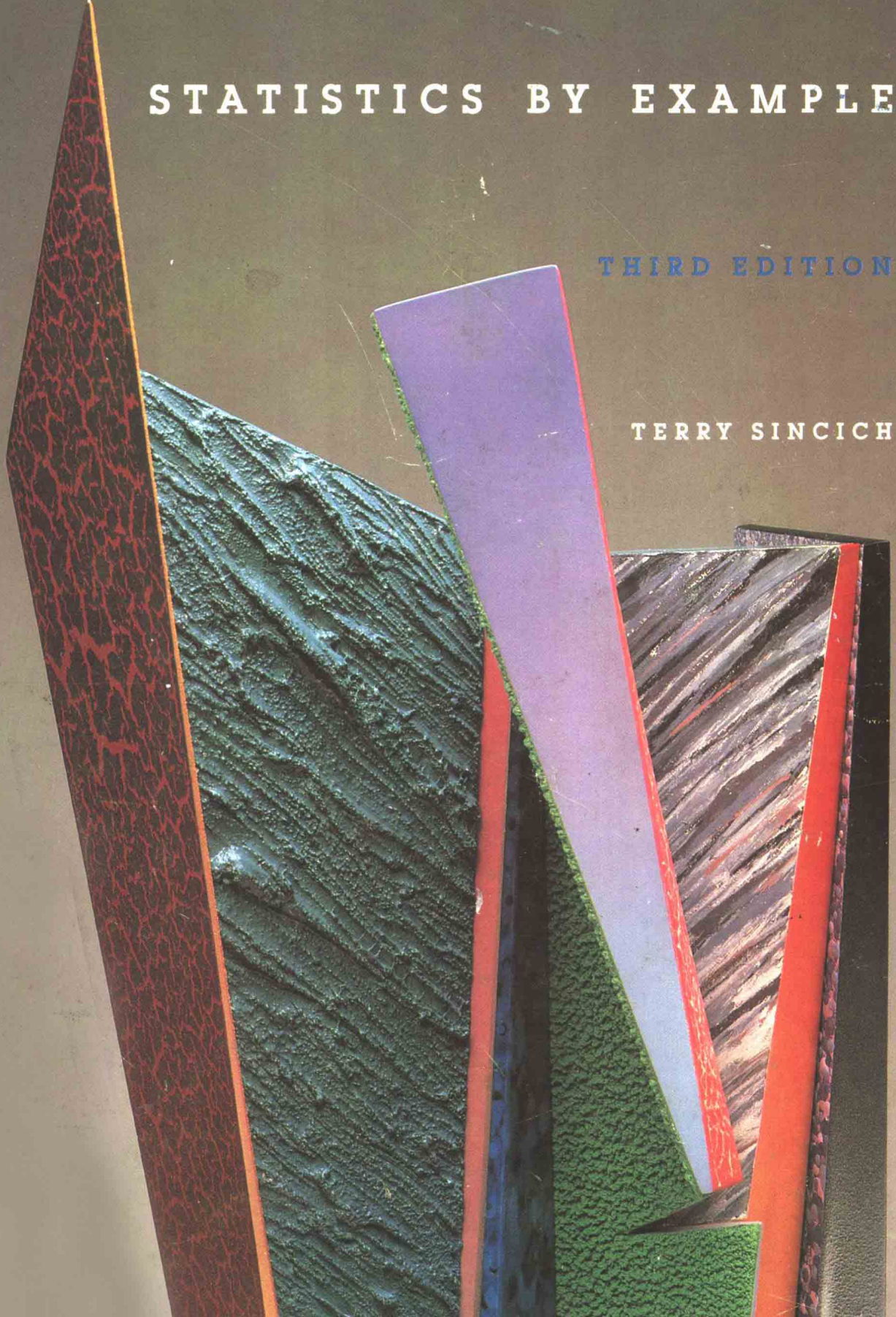
STATISTICS BY EXAMPLE

THIRD EDITION

TERRY SINCICH



THIRD EDITION

Statistics by Example

Terry Sincich
University of Florida

Dellen Publishing Company
San Francisco

Collier Macmillan Publishers
London

divisions of Macmillan, Inc.

On the cover: "Back Talk" by Stuart Lehrman, 1985; polychrome wood, 34" × 15" × 12". Stuart Lehrman combines sculpture with painting in works comprised of vibrantly colored and richly textured three-dimensional planes. Lehrman's work is included in many public and private collections in the United States. Lehrman is represented by Acme Art in San Francisco.

Printed in the United States of America

Permissions: Dellen Publishing Company
400 Pacific Avenue
San Francisco, California 94133

Orders: Dellen Publishing Company
c/o Macmillan Publishing Company
Front and Brown Streets
Riverside, New Jersey 08075

Collier Macmillan Canada, Inc.

Library of Congress Cataloging in Publication Data

Sincich, Terry.
Statistics by example.

Includes index.
1. Statistics. I. Title.
QA276.12.S56 1987 519.5 86–13569

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

ISBN 0-02-411031-0

PREFACE

This introductory college statistics text is designed for students who have only a high school background in mathematics as a prerequisite. It differs from most other texts in two ways:

1. **Real data sets** Explanations of basic statistical concepts and methodology are based on and motivated by the use of real data sets.
2. **Teaching by example** Concepts and statistical methods are explained in examples. These examples arise as questions posed about the data sets.

We think that this practical orientation helps the student to relate statistics to real-life problems and, hopefully, will develop a pattern of thought that will persist after the student enters the business world.

The text contains four data sets; the first two are heavily used as instructional vehicles. These data sets are:

Appendix A The set of actual starting salaries, majors, and colleges of 1,262 University of Florida graduates during the period from December 1981 to August 1983.

Appendix B The starting salaries (extracted from Appendix A) of bachelor's degree graduates of the Colleges of Business Administration, Education, Engineering, Liberal Arts, and Sciences.

Appendix C Supermarket customer checkout times for mechanical and automated checkers.

Appendix D Length, weight, and DDT measurements for various species of fish collected from the Tennessee River, Alabama, and its creek tributaries.

Although all of the data sets are used to develop the notion of a population and a sample, the starting salaries of college graduates (Appendices A and B) are used to demonstrate the need for data description, to develop the notion of a sampling distribution, and to motivate the inferential methods commonly studied in an introductory statistics course.

In addition to teaching via data sets and by example, this edition retains the following features of the first two editions:

1. **Case studies** At least three case studies that detail specific current events are used at the end of each chapter to pose questions for the student. These case studies, extracted from news articles and journals, demonstrate to the student the relevance of statistics to the solution of current practical problems. Answers to most case studies are included at the end of the text. (Refer to the Table of Contents for a list of the case studies.)
2. **Examples** The text, as its name implies, employs the "teaching by example" method. Each section contains several worked examples to demonstrate how to solve various types of statistical problems encountered in the real world.

3. **Key concepts highlighted** Throughout the text, key concepts are highlighted through the use of screened boxes.
   a. **Definitions** are boxed.
   b. **Steps** for constructing bar graphs, performing statistical calculations, and conducting statistical tests are listed and boxed for each procedure.
   c. **Key words**, which must be added to a student's vocabulary, are listed and boxed at the end of each chapter.
   d. **Warnings**, indicating situations where a student might misuse a statistical technique, are presented in boxed form. The student is directed to specific alternative methods.
4. **Probability and blackjack** The basic concepts of probability and their relation to statistical inference are presented in an easy-to-understand manner and are developed around the game of blackjack. Problem solving for the sake of problem solving is avoided.
5. **Many exercises** Since most students learn best by doing, the text contains a large number (over 1,000) of exercises. The answers for most are included at the end of the text. Each chapter contains exercises at the end of each section and a set of supplementary exercises at the chapter's end. The exercises are of two types:
   a. **Learning the Mechanics** These exercises are intended to be straightforward applications of the new concepts presented in the section. They are introduced in a few words and are unhampered by a barrage of background information designed to make them "practical," but which often detracts from instructional objectives. Thus, with a minimum of labor, the student can recheck his or her ability to comprehend a concept or definition.
   b. **Applying the Concepts** The mechanical exercises are followed by realistic exercises that allow the student to see applications of statistics to the solutions of a variety of real-world problems. Many of these exercises contain data extracted from newspaper articles, magazines, and journals. Once the mechanics are mastered, these exercises develop students' skills at comprehending realistic problems that describe situations to which the techniques may be applied.
6. **Computer printouts** The use of computer program packages is introduced in the presentation of the analysis of variance (Chapter 11), simple linear regression analysis (Chapter 12), multiple regression analysis (Chapter 13), and contingency table analysis (Chapter 14). The computer printouts for four different program packages, Minitab, SAS, SPSS^X, and BMDP, are presented for the analyses of identical sets of data.
7. **Integrating the data sets and the computer** The data sets can be entered into computer storage and can be accessed by students for sampling and statistical inference. The student (or instructor) can then access the data sets for the demonstration of statistical concepts and for realistic statistical exercises. For example, the data sets can be used by the instructor to illustrate the concept of a sampling distribution and the concepts of estimation and tests of hypotheses.

Although the scope of coverage remains the same, the third edition contains several substantial changes, additions, and enhancements:

1. **Case study section in all chapters** Each chapter now contains a case study section, in addition to the case studies presented at the end of the chapter. These new sections describe actual problems encountered in various fields of study and apply the statistical techniques learned in the chapter to solve the problems. Thus, each case study section gives an excellent example of the use of statistics in the real world.
2. **Chapter 7: Sampling and Sampling Distributions** A new section on the problems of nonresponse and invalid responses in survey sampling (Section 7.7) has been added.
3. **Chapter 11: Analysis of Variance** This chapter has been expanded to include randomized block designs. Three new sections have been added. The computing formulas for conducting an ANOVA with a randomized block design are given in Section 11.5, the formulas for confidence intervals and tests for differences between means in Section 11.6, and an example of a computer printout in Section 11.7.
4. **Chapter 15: Nonparametric Statistics** New sections on the Kruskal–Wallis H test for analyzing independent samples designs (Section 15.4) and the Friedman F_r test for analyzing randomized block designs (Section 15.5) have been added.
5. **More exercises with real data** Many new "real-world" exercises have been added to each chapter. These exercises, like the case studies, are extracted from news articles, magazines, and professional journals to give students the opportunity to apply their knowledge of statistics to current practical problems.

Numerous less obvious changes in details have been made throughout the text in response to suggestions by current users of the first two editions.

The text is also accompanied by the following supplementary material:

1. **Solutions manual** A student's exercise solutions manual presents the solutions for half the exercises contained in the text.
2. **Computer supplement** Easy-to-follow instructions on how to enter data into the computer and how to run most of the statistical procedures presented in the text are provided in a computer supplement. Each chapter in the supplement includes sample programs for four popular statistical program packages—SAS, SPSS^X, Minitab, and BMDP.
3. **Data sets available on diskette or tape** The data in the appendices are available on either an IBM PC diskette or a 7-track, nonlabelled magnetic computer tape.

I wish to acknowledge the many individuals who provided their invaluable assistance during the preparation of the original text and this revision. Their efforts are much appreciated. In particular, I thank the following for providing helpful suggestions and advice on the writing of the manuscript: Randal S. Beck, Millikin University;

Noni Bohonak, Dickinson College; David E. Booth, Millikin University; John S. Bowdidge, Southwest Missouri State University; Phillip G. Buckhiester, North Georgia College; Michael C. Burke, College of San Mateo; John Cameron, Rockhurst College; DeAnn Christianson, University of the Pacific; Jerry I. Goldman, DePaul University; William Hemmer, San Diego Mesa College; Geoffrey B. Holmewood, Hudson Valley Community College; Thomas B. Laase, University of Southern Colorado; James Lang, Valencia Junior College; Paul Lawrisuk, Moraine Valley Community College; James T. McClave, University of Florida; William Mendenhall, University of Florida; Franklin D. Rich, Sam Houston State University; and John B. Rushton, Metropolitan State College. Susan Reiland deserves special recognition for her excellent line-by-line reviews during the writing of the original manuscript and for managing the production of the second and third editions. I am very grateful to the following for providing the data sets and accompanying background information: Maurice Mayberry (Director, Career Resources Center, University of Florida), Jim Sullivan (Water and Air Research), Info Tech, and Venus Wong (who patiently spent long hours at the supermarket recording customer checkout times). Thanks are also due Faith Sincich for preparing answers to the exercises. I also wish to thank my typists, Brenda Dobson and Carol Springer, who did a remarkable job of converting my hand-scribbled notes into immaculate type.

Finally, I owe very special thanks to William Mendenhall and James T. McClave who, together, suggested the concept of this text and then provided me with the opportunity to write the manuscript. Without their guidance and encouragement this text and its subsequent revisions would never have been completed.

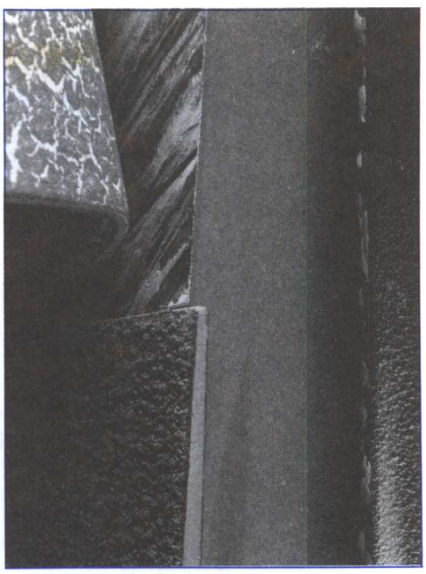
THIRD EDITION

Statistics by Example

CONTENTS

CHAPTER 1

Introduction



In one of the most famous experiments in psychology, Dr. Stanley Milgram (1974) studied the factors that determine the extent to which people obey authority—even if that authority is pushing them to do something they are against. In Milgram's study, subjects playing the role of "teachers" were told to give electric shocks (up to 450 volts) to "learners" who answered questions incorrectly. The object of the experiment was to see how many volts a subject would be willing to give before refusing to comply with the request.

Imagine the massive and impossible task of testing everyone in the country to see how they would respond to the shock experiment. And if we could test everyone, how would we determine what constitutes a "normal" reaction? This text describes the use of sampling and the science of statistics to solve many practical problems encountered in the real world. The case study in Section 1.3 discusses how Milgram measured people's obedience to authority and provides us with an opportunity to apply statistics to real-life data.

Contents