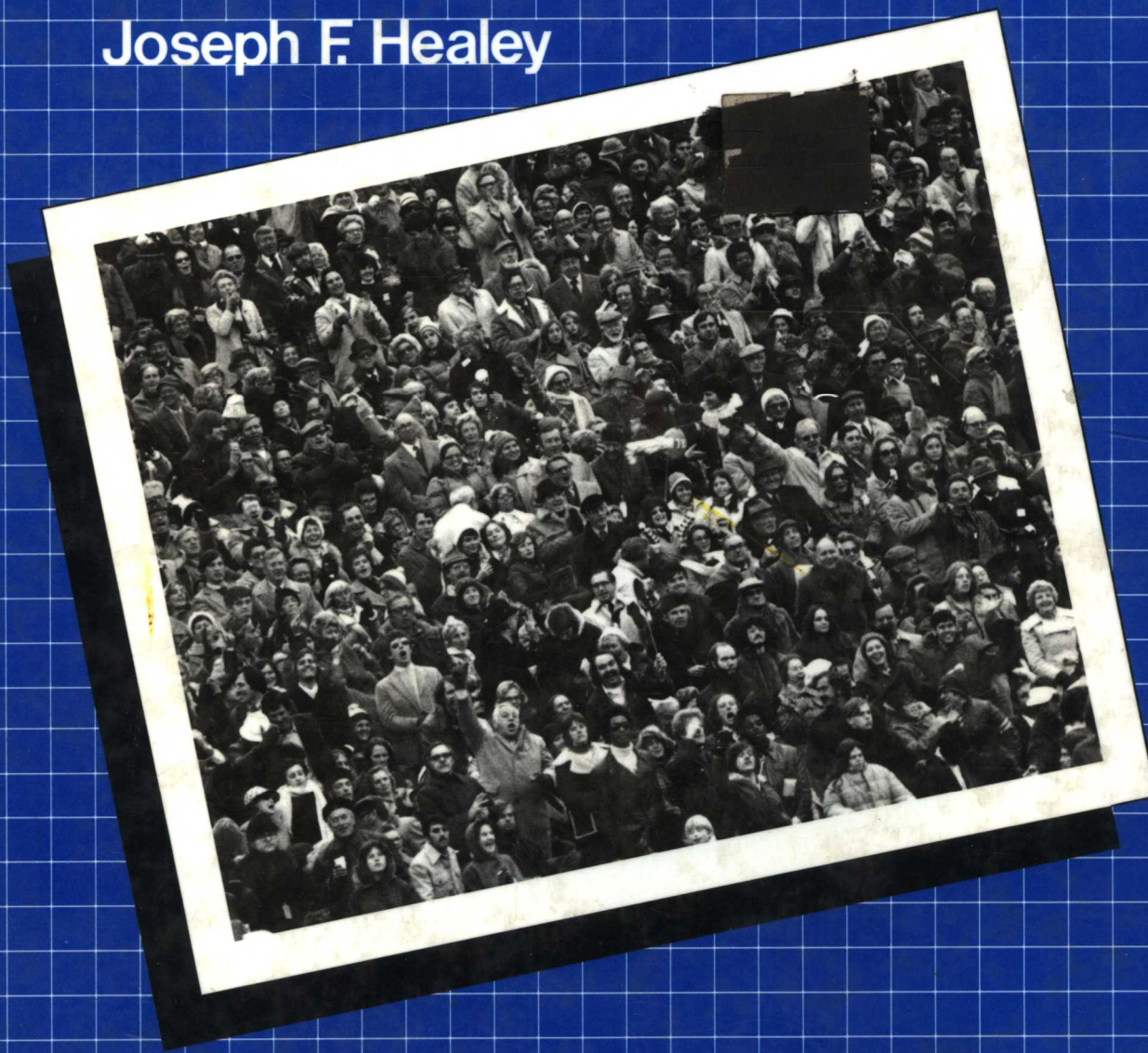
Statistics

A TOOL FOR SOCIAL RESEARCH

Joseph F. Healey



STATISTICS: A Tool for Social Research

Joseph F. Healey, Christopher Newport College

Wadsworth Publishing Company
Belmont, California
A Division of Wadsworth, Inc.

A study guide has been specially designed to help students master the concepts presented in this textbook. Order from your bookstore.

Sociology Editor: Sheryl Fullerton

Production Editor: Patricia Brewer

Managing Designer: Cynthia Bassett

Designer: Joseph di Chiarro

Copy Editor: Joan Pendleton

Technical Illustrator: Evanell Towne

Printed in the United States of America

3 4 5 6 7 8 9 10—88 87 86

Library of Congress Cataloging in Publication Data

Healey, Joseph F., 1945–
Statistics, a tool for social research.

Includes index.
1. Social sciences—Statistical methods. 2. Statistics.
I. Title.
HA29.H39 1984 519.5'024301 83-14691
ISBN 0-534-02985-X

STATISTICS: A Tool for Social Research

PREFACE

Teaching statistics to social science majors is, to say the least, a challenge. Students typically enter the classroom with a wide range of mathematical backgrounds and an equally diverse set of career goals. They are often genuinely puzzled about the relevance of statistics for their professional development and rather less than ecstatic about having the opportunity to gain some mastery of this subject matter. These elements of varied levels of preparation and need, combined with a certain reluctance to plunge head long into social statistics, make this course a pedagogical task of significant proportions.

This textbook was written to meet the challenge of introducing today's social science majors to statistical analysis while directly addressing these instructional problems. The text makes minimal assumptions about the mathematical background of the students (the ability to read a simple formula is sufficient preparation for virtually all of the material in the text), and a variety of special features have been integrated into the text to assist students to successfully analyze data. The text has been written especially for sociology and social work programs but is sufficiently flexible to be used in statistics courses in political science or in any program with a social science base and an applied focus (for example, public administration, criminal justice, urban studies, and gerontology).

The text is written at a level intermediate between the more rigorous and sophisticated texts on one hand and the mere "cookbook" on the other. That is, while I have not sacrificed comprehensive coverage or statistical correctness, the theoretical and mathematical explanations of why statistics "do what they do" are kept at an elementary level, as is appropriate in a first exposure to social statistics. For example, I do not treat formal probability theory per se. Rather, the background necessary for an understanding of inferential statistics is introduced, informally and intuitively, in Chapters 5 and 6 while considering the concepts of the normal curve and the sampling distribution. The text makes no claim that statistics are "fun" or that the material can be mastered without tears. At the same time, students are not overwhelmed with abstract proofs and mathematical theory, which at this level needlessly frustrate the learning experience.

My major goal is basic statistical literacy. The text is designed to provide a solid foundation in statistical analysis and to prepare students to be intelligent consumers of social research. More specifically, I believe that basic statistical literacy can be defined in terms of three interrelated qualities and, as a way of further describing the nature of this text, I would like to list each of these qualities and briefly summarize how the text is designed to develop them.

Computational Competence. At a minimum, students should emerge from their first course in statistics with the ability to perform elementary forms of data analysis; to execute a series of calculations and arrive at the correct answer. Since students in social science statistics courses frequently do not have strong quantitative backgrounds, I have included a number of features to help students cope with computations:

- *Step by step computational algorithms* are provided for each statistic.
- *Extensive problem sets* are provided at the end of each chapter. For the most part, these problems use fictitious data and are designed for relative ease of computation.
- *Solutions* to odd-numbered problems are provided so that students may check their answers.
- *An introduction to SPSS* is included in an appendix for instructors who wish to incorporate this computer application in the course.

An Appreciation of Statistics. A statistically literate person can do much more, of course, than merely calculate correct answers. Such a person understands the relevance of statistics for social research, can select an appropriate statistic for a given purpose and a given set of data, and can analyze and interpret the meaning of that statistic. This textbook begins to develop these qualities, within the constraints imposed by the introductory nature of the course, in the following ways:

- *The relevance of statistics.* Chapter 1 includes a discussion of the role of statistics in social research and stresses the usefulness of these techniques as ways of analyzing and manipulating data and answering research questions. Each example problem is framed in the context of a research problem. A question is posed and then, with the aid of a statistic, answered. The relevance of statistics for answering questions is thus stressed throughout the text. This central theme of usefulness is further reinforced by a series of boxes labeled "Applications," each of which illustrates some specific way statistics can be used to answer questions.

  Finally, the great majority of the end-of-chapter problems are labeled by the social science discipline or sub-discipline from which they are

drawn. Thus, students can more easily see the relevance of statistics to their own academic interests.

- *Selecting appropriate statistics.* A series of flowcharts are included to help students select appropriate statistics. These flowcharts have two components. Decision points are represented by diamonds and information by rectangles. The selection process is represented in general terms in the inside front covers and at the beginning of each part. Chapters begin with detailed flowcharts which, based on a consideration of the purpose of the analysis, the format of the data, and the level of measurement criterion, lead students to specific formulas or sections of the chapter. I have found these flowcharts very helpful in eliminating much of the confusion about "when to do what" that often afflicts beginning students.
- *Interpreting statistics.* After selecting and computing a statistic, students still face difficulties in understanding what the statistic means. The ability to interpret statistics can be developed only by exposure and experience. To provide exposure, I have been careful, in the example problems, always to express the meaning of the statistic in terms of the original research question. To provide experience, the end-of-chapter problems almost always call for an interpretation of the statistic calculated.

The Ability to Read the Professional Social Science Literature. The statistically literate person can comprehend and critically appreciate research reports written by others. The development of this quality is a particular problem at the introductory level because of the marked disparity between the concise language of the professional researcher and the rather wordy vocabulary of the classroom. The help bridge this gap, I have included a series of boxes labeled "Reading Statistics." These begin in Chapter 1 and appear every two or three chapters. In each box, I briefly describe the reporting style typically used for the statistic in question and try to alert students about what to expect when they approach the professional literature.

Additional Features. A number of other features make the text more meaningful for students and more useful for instructors:

- *Readability.* The writing style is informal and accessible to students without ignoring the traditional vocabulary of statistics. Problems and examples have been written to maximize student interest and to minimize the apparent dryness of the subject matter. For the more difficult material (such as hypothesis testing), students are first walked through an example problem before being confronted by formal terminology and concepts. Each chapter ends with a summary of major points and a glossary of important concepts. A glossary of symbols inside the back cover can be used for quick reference.

- *Organization and coverage.* The text is divided into four parts with most of the coverage devoted to univariate descriptive statistics, inferential statistics, and bivariate measures of association. The distinction between description and inference is introduced in the first chapter and maintained throughout the text.

  In selecting statistics for inclusion, I have tried to strike a balance between the essential concepts with which students must be familiar and the amount of material students can reasonably be expected to learn in their first (and perhaps only) statistics course, while bearing in mind that different instructors will naturally wish to stress different aspects of the subject. Thus, the text covers a full gamut of the usual statistics, with each chapter broken into subsections so that instructors may choose the particular statistics they wish to include.

- *Study Guide.* The Study Guide available for this text includes some extremely useful features. Each chapter includes a large set of review questions (multiple choice and fill in the blank), which students may use to check their comprehension of the material, and a large set of additional problems and exercises. Answers to all questions and solutions to all problems are provided.

  Each chapter contains calculator exercises designed to help students master the relevant computational routines. Also, two kinds of computer applications are included. Each chapter has exercises in the use of SPSS and, for instructors who wish to use microcomputers, the Study Guide has listings of BASIC programs for all major statistics. Finally, the Study Guide includes a social science data set for instructors who wish to have their students apply statistics to "real life" data.

- *Instructor's Manual.* The Instructor's Manual includes learning objectives, chapter summaries, a test item file of multiple choice questions, answers to even-numbered computational problems, and step-by-step solutions to selected problems.

Used with the supplements or by itself, *Statistics: A Tool for Social Research* is a comprehensive, flexible, and student-oriented textbook that provides a challenging first experience with social statistics.

Joseph F. Healey

ACKNOWLEDGMENTS

Many people have contributed to the development of this text. I am 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 reprint Appendixes B and C from their book *Statistical Tables for Biological, Agricultural and Medical Research* (6th edition, 1974).

I have had the pleasure of working with two different editors at Wadsworth and thank both Curt Peoples and Bill Oliver for their encouragement and assistance. Also, I would like to acknowledge the contributions of Bob Podstepny of Wadsworth during the early stages of this project and the excellent production editing of Pat Brewer.

Much of whatever integrity and quality this book has is a direct result of the contributions of colleagues who reviewed the manuscript during the various stages of writing. I have been consistently amazed by the sensitivity of these reviewers to the problems of writing an introductory statistics text and would like to thank David Bell, University of Houston; Kenneth J. Berry, Colorado State University; Clifford C. Clogg, Pennsylvania State University; Kathleen Crittendon, University of Illinois; Paula Dubeck, University of Cincinnati; Lawrence Felice, Baylor University; Gary D. Hampe, University of Wyoming; Anthony Andrew Hickey, George Mason University; Kichiro Iwamoto, University of Santa Clara; Jerry L. L. Miller, University of Arizona; Duane R. Monette, Northern Michigan University; Allen Rubin, University of Texas at Austin. Whatever failings the text retains are, of course, my responsibility and are probably the result of my occasional decisions not to follow the very sound advice of the reviewers.

I would also like to thank the instructors who made statistics understandable to me (especially Professors Satoshi Ito, Noelie Herzog, and Ed Erikson), all of my colleagues at Christopher Newport College for their support and encouragement (especially Professors Robert Durel, Ruth Kernodle, and William Winter), and especially my students for their patience and sound advice. For their help in the preparation and typing of the various drafts of the manuscript, I thank Wendy Turrissinni, Phyllis Snead Yeaw, Frank Marcucci, Evelyn Green, Christy Smith, and Patsy Trail. Special thanks to Debbie Kimsey for her assistance throughout the project and to Patricia Stuart, Marsha Butchko, and Elena Grose for their hard work at the final stages of preparation.

Last, I want to acknowledge the support of and dedicate this work to my entire family. I am fortunate to be a member of an extended family that is remarkable in many ways. Although I cannot list everyone (my family is also remarkably large), I would like to thank my mother, Alice Healey, and my aunt, Hazel Adams, for all they

have done for me. Finally, I owe most special thanks to my sons, Kevin and Christopher, and my wife, Judith Bessor Healey, for their patience and support throughout this project.

J. F. H.

BRIEF CONTENTS

DETAILED CONTENTS

chapter 1

INTRODUCTION

1.1 THE ROLE OF STATISTICS IN SCIENTIFIC INQUIRY

Students often approach their first course in statistics with questions about the value of the subject matter. What, after all, do numbers and statistics have to do with understanding people and society? In a sense, this entire book will attempt to answer this question, and the value of statistics will become clear as we move from chapter to chapter. For now, the importance of statistics can be demonstrated, in a preliminary way, by briefly reviewing the research process as it operates in the social sciences. These disciplines are scientific in the sense that social scientists attempt to verify their ideas and theories through research. Broadly conceived, research is any process by which information is systematically and carefully gathered for the purpose of answering questions, examining ideas, or testing theories. Research is a disciplined inquiry that can take numerous forms. Statistical analysis is relevant only for those research projects where the information collected is represented by numbers. Numerical information of this sort is called **data**, and the sole purpose of statistics is to manipulate and analyze data. **Statistics**, then, are a set of mathematical techniques used by social scientists to organize and manipulate data for the purpose of answering questions and testing theories.

What is so important about learning how to manipulate data? On one hand, let us admit that some of the most important and enlightening works in the social sciences do not utilize any statistical techniques. There is nothing magical about data and statistics. The mere presence of numbers guarantees nothing about the quality of a scientific inquiry. On the other hand, data can be the most trustworthy kind of information available to the researcher and, consequently, deserve special attention. Data that have been carefully collected and thoughtfully analyzed are the strongest, most objective foundations for building theory and enhancing understanding. Without a firm base in data, the social sciences would lose the right to the name "science" and would be of far less value to humanity.

Thus, the social sciences rely heavily on data-gathering for the advancement of knowledge. Let us be very clear about one point: it is never enough to merely gather data (or, for that matter, any kind of information). Even the most objective and carefully collected numerical information does not and