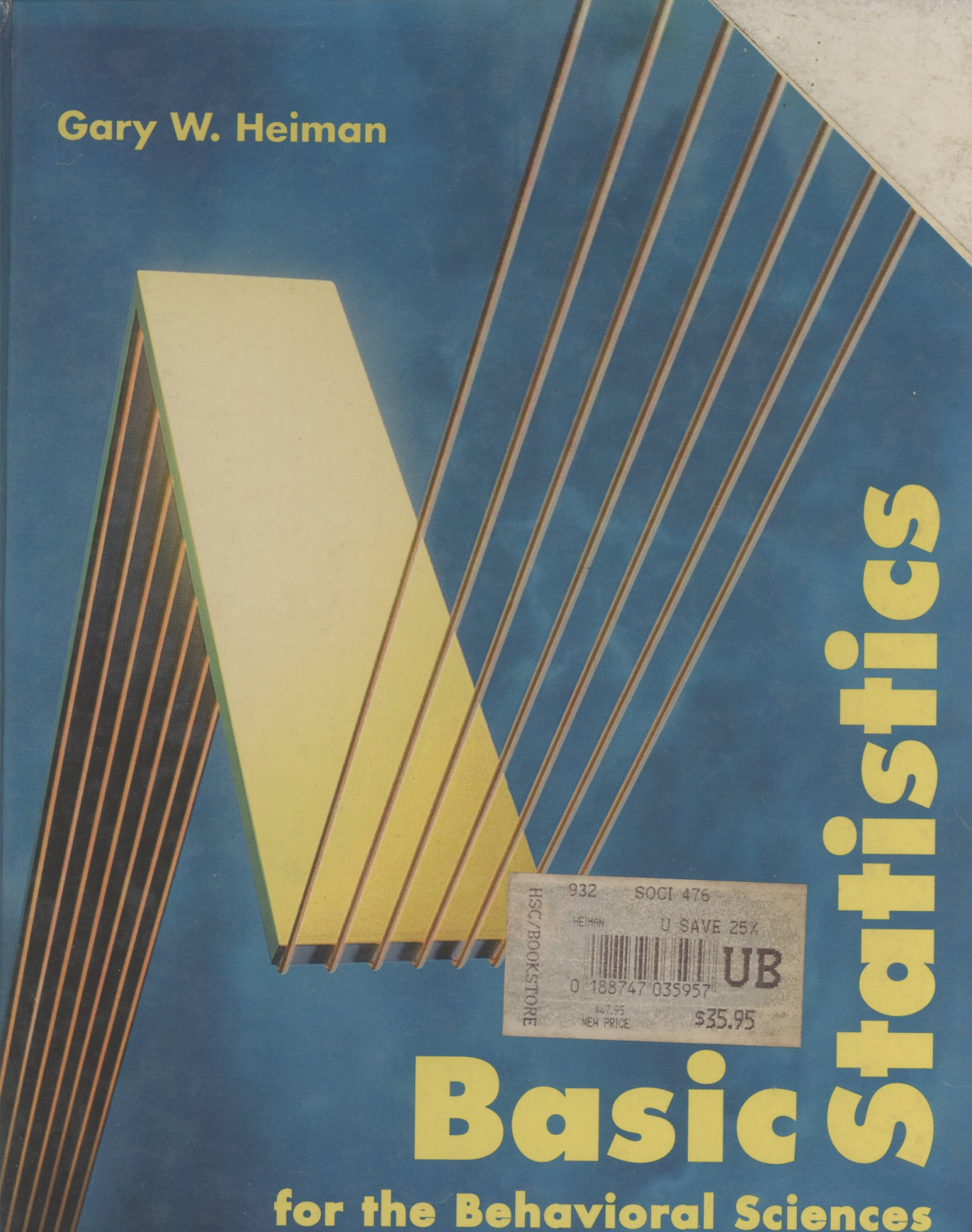
Gary W. Heiman

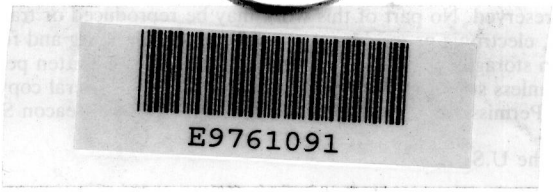


Basic Statistics

for the Behavioral Sciences

BASIC STATISTICS FOR THE BEHAVIORAL SCIENCES

Gary W. Heiman
State University of New York, College at Buffalo



HOUGHTON MIFFLIN COMPANY Boston Toronto
Dallas Geneva, Illinois Palo Alto Princeton, New Jersey

Sponsoring Editor: Michael DeRocco
Project Editor: Suzanne Morris
Production Coordinator: Frances Sharperson
Manufacturing Coordinator: Priscilla Bailey
Marketing Manager: Diane McOscar

Cover photograph by Michel Tcherevkoff/The Image Bank

Printed in the U.S.A.

Library of Congress Catalog Card Number: 91-71989

ISBN: 0-395-51546-7

ABCDEFGHIJ-D-987654321

BASIC STATISTICS FOR THE BEHAVIORAL SCIENCES

This book is dedicated to Jan, my most significant other.

Preface

Many of the undergraduates who enroll in my statistics course have a weak background in mathematics and some degree of "math phobia." By the end of the course, these same undergraduates must understand and be able to perform the descriptive and inferential statistical procedures that are commonly used in psychological and other behavioral research. In my fifteen-year search for a textbook that would help my students make this transition, I have never been fully satisfied.

A number of books dwell on the remarkable things statisticians do with statistics and say too little about the things *researchers* commonly do. Such books often present a catalog of procedures but do not explain the conceptual purposes of these procedures. Although students taught in this way can compute an answer on demand, they do not know why they should perform the procedure or what their answer reveals about the data.

My Objectives

I wanted my students to have a textbook that takes their needs more fully into account: a book that *explains*—clearly, patiently, and with an occasional touch of humor—the way a good teacher does. In striving to write such a book, I pursued five objectives.

1. Take a conceptual-intuitive approach To help students understand why as well as how procedures are performed, the text emphasizes the context in

which statistics are used to make sense out of data. Each procedure is introduced in the context of a simple study with readily understandable goals. I first focus on the purpose of research that examines relationships between variables, then delineate the procedures that describe and infer such relationships, and finally return to the conceptual purpose. Avoiding an abstract theoretical presentation, I instead provide students with simplified ways to think about statistical concepts and to see how these concepts translate into practical procedures for answering practical questions.

2. Present statistics within an understandable research context I assume that students have not taken a research methods course. For this reason I have created the research examples in this book rather than drawing them from the professional literature. Understanding examples from the literature typically requires a level of methodological sophistication that introductory statistics students do not possess, and the result is that the student is distracted from the illustrative purpose of the example. The text does contain the basic principles and terminology of research and is intended to help prepare students for a subsequent course in research methods.

3. Deal directly and positively with student weaknesses in mathematics The text presents no formulas or statistical statements without explanation. Formulas are introduced in terms of what they accomplish, and an example of each is worked out completely, step by step. To further reduce the apparent complexity of statistics, I have stressed the similarities among different procedures, showing how, despite slight variations in computations, they have similar components and answer similar questions.

4. Introduce new terms and concepts in an integrated way I have strived to tie each new concept and procedure to previous material, briefly reviewing that material in every possible instance. Difficult concepts are presented in small chunks, which are then built into a foundation and later elaborated on. My guiding rule was, to paraphrase, "We will serve no statistic before its time."

5. Create a text that students will enjoy as well as learn from To make the text readable and engaging, I have drawn on the many lessons I have learned from my students over the years. I repeatedly point out the everyday usefulness of statistics. I have tried to convey my own excitement about statistics and to dispel the notion that statistics (and statisticians) are boring. One can take a discipline seriously yet still recognize its quirks and foibles and have fun with it.

Pedagogical Format and Features

A number of features have been built into the book to enhance its usefulness as both a tool for study and a reference.

Conceptual and procedural questions, as well as computational problems, are provided at the end of each chapter. Both the final and the intermediate answers for all odd-numbered questions are given in Appendix E. Answers to the even-numbered questions appear in the Instructor's Resource Manual, which accompanies the text.

Other features include the following:

1. Graphs and diagrams are thoroughly explained in captions and fully integrated into the discussion.
2. New statistical notations are introduced at the beginning of the chapter in which they are used, not before.
3. Each definition is highlighted and is presented in clear and concise terms. Many mnemonics and analogies are included to promote retention and understanding.
4. Each important procedural point is emphasized by a "STAT ALERT." This is a summary reminder, set off from the text, about the calculation or interpretation of a statistic.
5. Every chapter summary provides a substantive review of the material, not merely a list of the topics covered.
6. For quick reference, formulas are listed at the end of each chapter.
7. A glossary of terms is included at the end of the book, and a glossary of symbols appears on the inside covers of the text.

Organization

The text is divided into four parts. In Part 1, *Getting Started,* Chapter 1 serves as a brief preface for the student and reviews basic math and graphing techniques. Chapter 2 then introduces the terminology, logic, and goals of statistics within the context of behavioral research. Chapters 3 through 6 make up Part 2, *Descriptive Statistics* (along with a discussion of linear interpolation in Appendix A). Part 3, *Describing Relationships,* consists of Chapters 7 and 8, in which correlation and regression are introduced as descriptive procedures, with emphasis on interpreting the correlation coefficient and the variance accounted for. (The point-biserial correlation is included to provide a bridge to measures of effect size in later chapters.)

Much of the text is organized around the bane of introductory students: inferential statistics. Extensive groundwork is laid in the chapters on descriptive statistics, with strong emphasis on understanding the proportion of the area

under the normal curve. Along with *z*-scores, Chapter 6 introduces the description of sample means using a sampling distribution.

Part 4, *Inferential Statistics,* begins with Chapter 9, which introduces probability. (Such topics as the additive and multiplicative rules and binomial expansion are presented in Appendix B.) The focus is on using the normal curve to compute probability, with the goal of making decisions about the representativeness of sample means. In Chapter 10, hypothesis testing is formalized using the *z*-test.

Chapter 11 presents the single-sample *t*-test, the confidence interval for a single mean, and significance testing of correlation coefficients. Chapter 12 covers two-sample *t*-tests, confidence intervals for the difference between the means and for the mean of differences, and effect size. Chapter 13 introduces the one-way, between-subjects ANOVA, including *post hoc* tests for equal and unequal *n*'s, eta squared, and omega squared. The F_{max} test is presented briefly. (The one-way repeated measures ANOVA is described in Appendix C.) Chapter 14 deals with the two-way between-subjects ANOVA, *post hoc* tests for main effects and for unconfounded comparisons in an interaction, as well as graphing and interpreting interactions. Chapter 15 covers the one-way and two-way chi square, as well as the nonparametric versions of all previous parametric tests (with appropriate *post hoc* tests and measures of effect size).

The text strives to teach students how to interpret their data—not just to report that a result is significant. Thus, I have emphasized such topics as plotting and interpreting graphs and understanding the relationships demonstrated by research. I've also included practical discussions of power and measures of effect size. These discussions occur at the end of a section or chapter so that instructors wishing to skip these topics can easily do so.

Software

A data-analysis computer program called HMSTAT has been custom-tailored to this text by Dr. David Abbott of the University of Central Florida in Orlando. The program is packaged free with the student text. This menu-driven program can accept and store data, perform all the procedures discussed in the text, and be operated by students with a minimal computer background. The program is integrated with the text only through the final section in each chapter, "Using the Computer." Otherwise, use of the software is entirely optional.

Other Ancillaries

Additional practice problems are available in the Student Workbook and Study Guide, which I co-wrote with my colleague Dr. Deborah Kohl. Each chapter contains a review of objectives, terms, and formulas, a programmed review, and conceptual/computational problems (answers are included). Each chapter

also has a set of test-like questions for which answers are provided in the Instructor's Resource Manual. An additional chapter review facilitates student integration of the entire course.

The Instructor's Resource Manual, by Dr. David Chattin of St. Joseph College, Indiana, contains approximately 750 test items and problems, as well as suggestions for classroom activities and discussion. The test items are available on computer disk in an ASCII-file format.

Acknowledgments

I gratefully acknowledge the help and support of many professionals at Houghton Mifflin Company. In particular, I want to thank my outstanding sponsoring editor, Michael DeRocco, who provided invaluable guidance and support, and Karen Donovan, who as developmental editor served above and beyond any reasonable expectation. I am grateful to Suzanne Morris and Sally Lifland for seeing the book through production and to Catherine Johnson for her excellent interior design. Dr. Kathleen Donovan of the University of Central Oklahoma checked the entire manuscript for accuracy, and I thank her for her diligence. Any errors that might remain, however, are mine and mine alone. Finally, I am grateful to the following reviewers who, in evaluating all or parts of this manuscript at one stage or another, provided invaluable feedback.

William Addison, Eastern Illinois University
Andrew A. Beveridge, Queens College, City University of New York
Elliott Bonem, Eastern Michigan University
David Chattin, St. Joseph College, Indiana
Charles L. Dufour, Colby College
William A. Frederickson, University of Central Oklahoma
Philip S. Gallo, Jr., San Diego State University
Jane A. Halpert, DePaul University
Milton H. Hodge, University of Georgia
Eric S. Knowles, Ohio State University
Suzanne Mannes, University of Delaware
Richard Rogers, Georgia Southern College
Glenn S. Sanders, State University of New York, Albany
Billy L. Smith, University of Central Arkansas
James L. Walker, Jr., Lamar University
Christine B. Ziegler, Kennesaw College

Gary W. Heiman

Contents

2

Statistics and the Research Process 17

PART 2

DESCRIPTIVE STATISTICS: DESCRIBING SAMPLES AND POPULATIONS

3

Frequency Distributions and Percentiles 42

4

Measures of Central Tendency: The Mean, Median, and Mode 76

5

Measures of Variability: Range, Variance, and Standard Deviation 105