

Gary W. Heiman



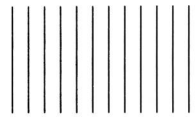
Basic Statistics

for the Behavioral Sciences

0212

9761091

H467



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

Free Copy
Not to Be Sold

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

Copyright © 1992 by Houghton Mifflin Company.

All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system without the prior written permission of Houghton Mifflin Company unless such copying is expressly permitted by federal copyright law. Address inquiries to College Permissions, Houghton Mifflin Company, One Beacon Street, Boston, MA 02108

Printed in the U.S.A.

Library of Congress Catalog Card Number: 91-71989

ISBN: 0-395-51546-7

ABCDEFGHIJ-D-987654321

Free Copy
Not to Be Sold



This book was originally distributed by the publisher as a sample copy for academic review. This book was subsequently purchased by a book dealer and classified as **"USED"** which allows you, the student, to buy this book at a substantial savings. *All chapters and pages are included.*

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

QA

519.5

***Free Copy
Not to Be Sold***

此为试读, 需要完整PDF请访问: www.ertongbook.com



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

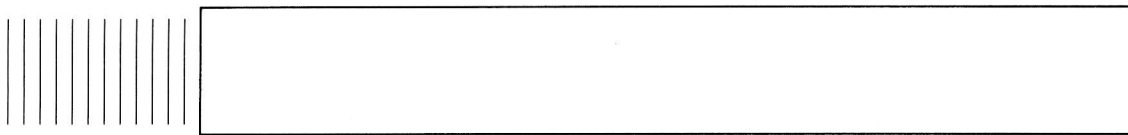
Glenn S. Sanders, State University of New York, Albany

Billy L. Smith, University of Central Arkansas

James L. Walker, Jr., Lamar University

Christine B. Ziegler, Kennesaw College

Gary W. Heiman



Contents

Preface

xxiii

PART 1

GETTING STARTED

1

Introduction 2

Some Answers to Your Questions and Concerns About Statistics	2
What Are Statistics? 2	
What Do Psychologists Do with Statistics? 2	
But I'm Not Interested in Research; I Just Want to Help People! 3	
But I Don't Know Anything About Research! 3	
What If I'm Not Very Good at Statistics? 3	
But Statistics Aren't Written in English! 4	
What If I'm Not Very Good at Math? 4	
So All I Have to Do Is Learn How to Compute the Answers? 4	
What Should Be My Approach? 5	
All Right, So How Do I Learn Statistics? 5	
What Should I Know About This Book? 5	

Free Copy
Not to Be Sold

vii

Review of Mathematics Used in Statistics	6
Basic Statistical Notation 7	
<i>Using symbols for mathematical operations 7</i>	
<i>Determining the order of mathematical operations 7</i>	
<i>Solving equations to find the answer 8</i>	
Rounding 9	
Transformations 10	
<i>Proportions 10</i>	
<i>Percents 10</i>	
Creating Graphs 11	
Finally	13
Chapter Summary	14
Practice Problems	15

2

Statistics and the Research Process 17

The Logic of Scientific Research	17
Variables 17	
Relationships Between Variables 18	
<i>Degree of association in a relationship 19</i>	
<i>No relationship 21</i>	
<i>Graphing relationships 21</i>	
<i>Using relationships to discover laws 22</i>	
Samples and Populations	23
Representativeness of a Sample 24	
Random Sampling 25	
Research Designs	26
Experiments 26	
<i>The independent variable 27</i>	
<i>The dependent variable 28</i>	
<i>Drawing conclusions about the population 29</i>	
<i>The problem of causality 29</i>	
Correlational Studies 30	
Using Statistical Procedures to Analyze Data	31
Descriptive Statistics 32	
Inferential Statistics 33	
Statistics and Parameters 34	
Types of Variables	34
The Four Types of Measurement Scales 35	
Discrete and Continuous Scales 36	
Finally	37
Chapter Summary	37
Using the Computer	39
Practice Problems	39

Free Copy
Not to Be Sold

PART 2

DESCRIPTIVE STATISTICS: DESCRIBING SAMPLES AND POPULATIONS

3

Frequency Distributions and Percentiles 42

More Statistical Notation	42
Creating Simple Frequency Distributions	43
Presenting Simple Frequency in a Table	43
Graphing a Simple Frequency Distribution	45
<i>Bar graphs</i> 45 <i>Histograms</i> 45 <i>Frequency polygons</i> 47	
Types of Simple Frequency Distributions	48
The Normal Distribution	48
<i>Overlapping normal distributions</i> 51 <i>Approximations to the normal distribution</i> 51 <i>The normal curve model</i> 52	
Other Common Frequency Polygons	52
<i>Skewed distributions</i> 53 <i>Bimodal and rectangular distributions</i> 53	
<i>Distributions of real data versus ideal distributions</i> 54	
Creating Relative Frequency Distributions	55
Presenting Relative Frequency in a Table	56
Graphing a Relative Frequency Distribution	57
Finding Relative Frequency Using the Normal Curve	57
Creating Cumulative Frequency Distributions	59
Presenting Cumulative Frequencies in a Table	59
Graphing a Cumulative Frequency Distribution	60
Creating Grouped Frequency Distributions	60
Real Versus Apparent Limits	63
Graphing Grouped Distributions	64
<i>Graphing grouped simple and relative frequency distributions</i> 64	
<i>Graphing grouped cumulative frequency distributions</i> 64	
Computing Percentile	66
Determining the Score at a Given Percentile	66
Finding a Percentile for a Given Score	69
Finding Percentile Using the Area Under the Normal Curve	70
Finally	71
Chapter Summary	72
Using the Computer	74
Practice Problems	74
Summary of Formulas	75

Free Copy
Not to Be Sold

4

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

More Statistical Notation	76
What Is Central Tendency?	77
The Mode	78
Uses of the Mode	79
Problems with the Mode	80
The Median	81
Uses of the Median	83
Problems with the Median	83
The Mean	83
Uses of the Mean	85
Problems with the Mean	86
Transformations and the Mean	88
Deviations Around the Mean	88
Using the Mean to Predict Individual Scores	90
Using Deviations to Indicate a Score's Location in a Distribution	91
Using the Sample Mean to Describe the Population Mean	92
Summarizing the Results of Experiments	94
Summarizing a Relationship Using Measures of Central Tendency	95
Graphing the Results of an Experiment	96
<i>Line graphs</i>	96
<i>Bar graphs</i>	98
Describing the Relationship in the Population	99
Finally	101
Chapter Summary	101
Using the Computer	103
Practice Problems	103
Summary of Formulas	104

5

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

More Statistical Notation	105
The Concept of Variability	106

*Free Copy
Not to Be Sold*

The Range	108
The Semi-interquartile Range	109
Understanding Variance and Standard Deviation	110
Describing the Sample Variance	111
Computational Formula for the Sample Variance	113
Interpreting Variance	115
Describing the Sample Standard Deviation	115
Computational Formula for the Sample Standard Deviation	116
Interpreting the Standard Deviation	117
Applying the Standard Deviation to the Normal Curve	117
Avoiding Some Common Errors in Computing S_x	120
Mathematical Constants and the Standard Deviation	121
The Population Variance and the Population Standard Deviation	121
Estimating the Population Variance and Population Standard Deviation	122
<i>Computational formula for the estimated population variance</i>	125
<i>Computational formula for the estimated population standard deviation</i>	126
Interpreting the Estimated Population Variance and Standard Deviation	127
Variance Is the Error in Predictions	127
Estimating the Error in Predictions in the Population	128
Accounting for Variance	129
Finally	131
Chapter Summary	132
Using the Computer	134
Practice Problems	134
Summary of Formulas	135

6

z-Score Transformations and the Normal Curve Model 137

More Statistical Notation	138
Understanding z-Scores	138
Describing a Score's Relative Location as a z-Score	140
Computing z-Scores	141
Computing a Raw Score When z Is Known	142
How Variability Influences z-Scores	143

*Free Copy
Not to Be Sold*

Interpreting z-Scores: The z-Distribution	144
Characteristics of the z-Distribution	146
Transforming z-Scores	146
Using the z-Distribution to Compare Different Distributions	147
Plotting Different z-Distributions on the Same Graph	149
Comparing the Relative Frequency on Different z-Distributions	150
Using the z-Distribution to Determine the Relative Frequency of Raw Scores	150
The Area Under the Standard Normal Curve	151
Applying the Standard Normal Curve Model	152
<i>Finding percentile rank for a raw score</i>	154
<i>Finding a raw score at a given percentile rank</i>	155
<i>Using the z-table</i>	155
Using the Standard Normal Curve Model to Define Psychological Attributes	158
Using the Standard Normal Curve Model to Describe Samples	158
The Sampling Distribution of Means	159
The Standard Error of the Mean	161
Using the Sampling Distribution to Determine Relative Frequency of Sample Means	162
Finally	165
Chapter Summary	166
Using the Computer	167
Practice Problems	168
Summary of Formulas	169

PART 3

DESCRIBING RELATIONSHIPS

7

Describing Relationships Using Correlations 172

More Statistical Notation	172
Understanding Correlational Research	173
Correlation and Causality	174
Distinguishing Characteristics of Correlation	175
Plotting Correlational Data: The Scatterplot	176

Free Copy
Not to Be Sold

The Correlation Coefficient	178
Types of Relationships	178
Linear Relationships	179
Nonlinear Relationships	180
How the Correlation Coefficient Describes the Type of Relationship	181
Strength of the Relationship	182
Perfect Association	183
Intermediate Association	184
Zero Association	186
Computing the Correlation Coefficient	187
The Pearson Correlation Coefficient	188
<i>Computing the Pearson correlation coefficient</i>	188
The Spearman Rank-Order Correlation Coefficient	190
<i>Computing the Spearman correlation coefficient</i>	192
<i>Tied ranks</i>	193
The Point-Biserial Correlation Coefficient	193
<i>Computing the point-biserial correlation coefficient</i>	194
The Restriction of Range Problem	196
Correlations in the Population	197
Finally	198
Chapter Summary	199
Using the Computer	200
Practice Problems	201
Summary of Formulas	204

8

Linear Regression and Predicting Variability 205

More Statistical Notation	205
Understanding Linear Regression	206
Summarizing the Scatterplot Using the Regression Line	206
Predicting Scores Using the Regression Line	208
The Linear Regression Equation	209
Computing the Slope	212
Computing the Y-Intercept	213
Describing the Linear Regression Equation	213
Plotting the Regression Line	213
Using the Regression Equation to Predict Y Scores	214

Free Copy
Not to Be Sold