

Hinkle Wiersma Jurs

BASIC BEHAVIORAL STATISTICS

Basic Behavioral Statistics

Dennis E. Hinkle
Virginia Polytechnic Institute
and State University

William Wiersma
The University of Toledo

Stephen G. Jurs
The University of Toledo

Houghton Mifflin Company Boston
Dallas Geneva, Illinois Hopewell, New Jersey Palo Alto London

To
Mary, Chas, and Anya
and
Joan, Lisa, and Susan
and
Sara, Peter, and Andrew

Ceramic art on cover and chapter openers by Kenneth Goldstrom.

Copyright © 1982 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, except as may be expressly permitted by the 1976 Copyright Act or in writing by the publisher. Requests for permission should be addressed in writing to Houghton Mifflin Company, One Beacon Street, Boston, Massachusetts 02108.

Library of Congress Catalog Card Number: 81-81700

ISBN: 0-395-31729-0

Printed in the U.S.A.

Foreword to the Student

By Sheila Tobias

author of *Overcoming Math Anxiety* (Houghton Mifflin, 1980)

For some very unusual people mathematics is a gift. For the rest of us it is a collection of learnable (and easily forgettable) skills. Hence it may be wise, before you undertake to learn the material covered in this text, to do some systematic review of the algebra and precalculus courses you may have taken, along with a group of supportive and like-minded friends. Such re-exposure will remind you of mathematical notation and of some of the basic algebra you once knew. And it may help you “hit the ground running” when you start statistics.

A more problematical aspect of mathematics, somewhat harder to deal with, is the built-in prejudice some of us bring to mathematics and math-related fields of study. We call this phenomenon “math anxiety” or “math avoidance,” and in numerous intervention programs around the country, students are being urged—before they confront their required statistics course—to undertake some “math anxiety reduction.” Such a process involves self-examination of your own, personal math autobiography, some self-monitoring of the ways you may be defeating yourself when doing mathematics, and some study of mathematics in a tension-free and pressure-free environment.

Another technique that is used for reducing math anxiety is assertiveness training. Some of us feel embarrassed when we don’t “get” a mathematical idea the first time it is presented, and consequently we don’t ask questions in class. Assertiveness training teaches students how to ask their instructors for the help they need. Sometimes the text—even one as carefully prepared as this one—may not answer a question floating around in your mind. Note such questions and ask for clarification. If nec-

essary, write your questions down and give them to your instructor. Being cumulative, your statistics course will build continuously on everything that has gone before. A missed concept or a misunderstanding is like a dropped stitch; it is all the harder to correct later on.

A third helpful technique is to work with other people. Your instructor may not permit you to take group tests (though some teachers are beginning to allow this in order to reduce tension and competitiveness among students), but group study is often worthwhile. Explaining concepts to one another enhances and deepens your own learning, and the group experience relieves some of the uncertainty of doing mathematics in isolation.

Above all, don't be demoralized by your mistakes. In a session held for math-anxious adults, a graduate student in mathematics once agreed to answer general questions from the group. An intrepid woman ventured this question: "How do you feel when you make a mistake?" After a short pause, the graduate student answered, "I find my mistakes interesting and my confusions even more so." The group gasped. They had come to believe that mistakes were like blemishes, emblems of their stupidity. The graduate student had pointed out, correctly, that mistakes are like windows into our minds.

Examine your mistakes. Are they the result of careless error—something you wouldn't do again—or do they indicate a missing skill? Did you mislearn an operation that worked for awhile on simpler data but will not work on the new material? Is the problem in the mathematics? In your comprehension of the verbal information presented? In the notation? (You will be learning more new notation in your statistics course than you learned in your algebra courses in high school. Some of this notation will be different from and contradictory to what you learned before.) Or is the problem in visualizing the operation when it is expressed on a graph?

Obviously, the better you analyze your errors (alone or with the help of your tutor or instructor), the less likely it is that you will make them in the future.

Finally, be flexible. Even mathematicians may modify a problem they are working on to get a better handle on it. They substitute smaller numbers or replace a long series with a shorter one, just to get an idea of the parameters of the problem and to test their methods. Guessing is O.K., so long as you find a good way to test your hunches. (Sometimes the way you test your hunch will give you a good clue to the formula you are seeking.) And use your intuition. Math-anxious people trust their intuition far less than successful learners of math and statistics.

This textbook has been skillfully prepared and carefully reviewed by many experienced teachers of statistics, with the aim of making the presentation as clear as possible. But this does not mean you will learn everything you need to know the first time around. "Reading" mathematics or statistics is very different from reading social science or general nonfiction. Read with a paper and a pencil; each statement is made once and you must master it before going on. Imagine examples, imagine counter-examples. Note any questions that arise and, if these are not answered by the time you are finished studying, give them to your instructor. Try to slow down your reading speed. Think of reading statistics as rather like "reading" a crossword puzzle. You don't just look at the cues for Down and Across; you work them out. And the same

is true for this textbook. Don't just glance at the words; work out the ideas and apply them.

If you have anxiety to deal with, try to make yourself relax. Writing down your "self-defeating self-talk" is one way to relax. Try to get more time to complete your tests if your instructor will allow it. And you may want to consult other books on math anxiety.

Consider ways to make your first course in statistics easier for you. You might audit the course for one semester before taking it for credit. This will give you an idea of the scope of the material and some advance warning about what the more difficult areas will be. You might even discover that the subject of statistics is both fascinating and empowering: The opportunity to take control of data, to find patterns in ostensibly unrelated phenomena, and to deal critically with other people's research will give you thinking strengths that no other course can offer.

Preface

This book was written primarily as a text for a one-term or semester course in statistics. However, it can serve as a supplement for related courses with a quantitative orientation, such as courses in measurement and research methodology. It is an introductory text and is designed to teach concepts and procedures that are fundamental for further study in statistics. The book can be used for either undergraduate or graduate courses found in behavioral sciences curricula such as psychology, education, and sociology.

Conceptual Approach The approach of the text is conceptual and nonmathematical in nature. Any study of statistics must include some consideration of formulas and computation, but we have included formulas only when necessary and then with an emphasis on understanding the formula, not a “cookbook” use of it. We have selected the simplest version of formulas in most cases, and we do not confuse the student with three or four equivalent formulas. Basic concepts of statistical reasoning and the underlying assumptions are presented. For undergraduate students and beginning graduate students (especially those who may take additional statistics courses) such concepts are necessary, even though they require only a limited mathematical background. Computation is illustrated through the examples that are introduced to develop understanding of the concepts.

Learning and Teaching Aids The book contains several features designed to enhance its use as a teaching tool and a learning device. Statements summarizing the

important points being made are displayed prominently throughout the chapters. Key concepts listed at the end of each chapter help the reader review the major concepts introduced in the chapter. Exercises are included for each chapter, and solutions to the exercises are provided in an appendix. Thus students can receive immediate feedback about their solutions. For students who use an electronic calculator to work problems, the key strokes for selected exercises have been included at the end of each chapter in a section identified by a calculator symbol, as on page 21. The use of pocket calculators in statistical analysis is described in Appendix A.

A *Student Workbook* is available and can be used effectively as a supplement to the text. It includes definitions of key concepts, exercises to reinforce these concepts, and additional exercises for further review.

The first chapter offers an optional review of elementary mathematics for those students with a very limited mathematics background. This review covers the basic operations of addition, subtraction, multiplication, and division, along with the real number system and elementary algebra. Notation is explained, and the Σ operator, so important in statistics, is introduced. But essentially, no mathematical background beyond elementary algebra is required for using the text effectively.

Coverage The book includes both descriptive and inferential statistics, the latter through analysis of variance. The basic descriptive statistics are presented in Chapters 2 through 7; included are graphs, measures of location, measures of central tendency, measures of variability, standard scores, correlation, and regression. The learning process is enhanced through the use of realistic examples and exercises. It should be noted that the concept of regression is considered in the descriptive sense—that is, in terms of the development of the regression equation, the use of this equation in the prediction process, and the definition of the standard error of prediction.

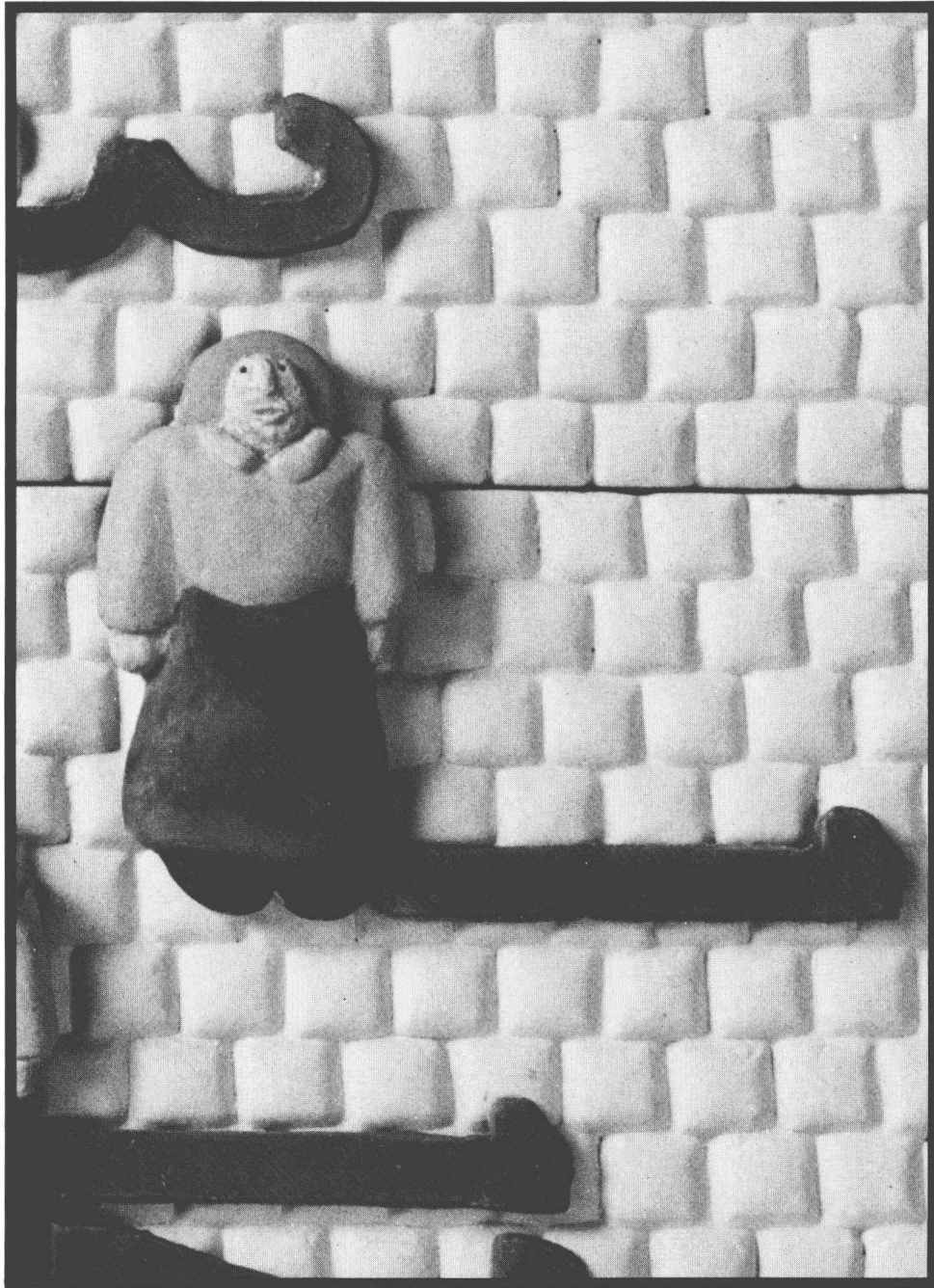
Inferential statistics are introduced via a discussion of the concept of probability and its relationship to sampling and sampling distributions. These basic concepts are then used in developing the logic underlying hypothesis testing and parameter estimation. Hypothesis testing and estimation are discussed for the one-sample case for the mean in Chapters 9 and 10 and then expanded to the two-sample case in Chapter 11. One-sample and two-sample cases for other parameters are considered in Chapter 12.

Chapter 13 is devoted to one-way analysis of variance (ANOVA). Conceptual development of the partitioning of the sum of squares and *post hoc* multiple comparison procedures receive special consideration. Repeated measures analysis of variance is also included. Two-factor analysis of variance is treated in Chapter 14 as the logical extension of one-way ANOVA. Selected nonparametric statistics are presented in Chapter 15; these procedures are the analogues to the procedures discussed in Chapters 9 through 13.

Acknowledgments We are 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 Tables III, IV, VII, and XXIII from their book *Statistical Tables for Biological, Agricultural and Medical Research* (6th edition, 1974).

We would like to acknowledge several people who have contributed to this effort. First of all we would like to thank the reviewers who gave us most constructive criticisms as well as suggestions that have enhanced the manuscript: William A. Frederickson, Central State University; John Harsh, University of Southern Mississippi; Donald V. Huard, Phoenix College; Charles R. Kessler, Los Angeles Pierce College; and Alice L. Palubinskas, Tufts University. Special thanks goes to J. R. Cox for his review of the manuscript and to L. E. Ewing for his assistance in developing the exercises and their solutions. Our most special acknowledgment is to our families, to whom we dedicate the book.

Dennis E. Hinkle
William Wiersma
Stephen G. Jurs



Contents

Foreword to the Student by *Sheila Tobias* xv

Preface xix

1

Introduction 1

The Meaning of Statistics 2

Basic Mathematics Needed for Statistics 2

Four Basic Arithmetic Operations 2

Arithmetic Operations with Real Numbers 3

Rounding Numbers 5

Postulates for Real Numbers 6

Special Indicators of Arithmetic Operations 7

The Summation Operator 9

Variables 11

Continuous and Discrete Variables 11

Independent and Dependent Variables 12

Measurement Scales 12

Nominal Scale 13

Ordinal Scale	14
Interval Scale	14
Ratio Scale	15
Populations and Samples	16
Descriptive and Inferential Statistics	16
Summary	17
Key Concepts	18
Exercises	18
Key Strokes for Selected Exercises	21

2

Frequency Distributions 23

Frequency Distributions—A Method for Organizing Data	24
Exact Limits of the Class Interval	26
Rules for Developing Class Intervals	27
Assumptions for Class Intervals	28
Graphing Data	29
The Nature of Graphs	30
Bar Graph	31
Data Curves	32
Scattergrams	32
Graphing Frequency Distributions	33
Histograms	34
Frequency Polygons	35
The Cumulative Frequency Distribution and Its Graph	36
Various Shapes of Frequency Distributions	37
Summary	40
Key Concepts	40
Exercises	41

3

Describing Distributions 47

Measures of Central Tendency	48
Mode	49
Median	50
Mean	52

Weighted Mean—The Mean of Combined Groups	54
Comparison of Mode, Median, and Mean	55
Measures of Variation	57
Range	59
Mean Deviation	59
Variance	61
Standard Deviation	63
Summary	64
Key Concepts	65
Exercises	65
Key Strokes for Selected Exercises	67

4

Describing Individual Scores 69

Percentiles	70
Percentile Rank	72
Use of Percentiles and Percentile Ranks	74
Standard Scores	75
Properties of Standard Scores	77
Transformed Standard Scores	78
Weighted Averages for Several Tests	80
Summary	81
Key Concepts	81
Exercises	81
Key Strokes for Selected Exercises	83

5

The Normal Curve 85

The Nature of the Normal Curve	86
The Standard Normal Curve	89
Using the Standard Normal Curve	90
Normalized Standard Scores	96
Summary	98
Key Concepts	99
Exercises	99
Key Strokes for Selected Exercises	100

6

Correlation 103

- The Concept of Correlation 105
 - Scattergrams 105
- The Pearson Product-Moment Correlation Coefficient 108
 - Computational Formula for the Correlation Coefficient 108
 - Interpreting the Correlation Coefficient 109
 - The Assumption of Linearity 109
- The Spearman Rho Correlation Coefficient 111
- Other Correlation Coefficients 113
- Factors that Affect Correlations 113
- Uses of the Correlation Coefficient 115
- Summary 116
- Key Concepts 116
- Exercises 117
- Key Strokes for Selected Exercises 119

7

Regression and Prediction 123

- Correlation and Prediction 124
- The Regression Line 125
 - Determining the Regression Line 127
- Errors in Prediction 130
 - The Standard Error of Estimate 131
- Predicted Scores and Probability Statements 132
- The Second Regression Line 134
- Correlation and Causation 135
- Summary 136
- Key Concepts 137
- Exercises 137
- Key Strokes for Selected Exercises 138

8

Probability 141

- Probability—A Definition 142
- Laws of Probability 143

Addition Rule for Mutually Exclusive Events	143
Addition Rule for Non-Mutually Exclusive Events	145
Multiplication Rule for Statistically Independent Events	146
Multiplication Rule for Nonindependent Events	147
Probability and the Underlying Distribution of All Possible Outcomes	148
The Binomial Distribution as an Underlying Distribution	149
The Normal Distribution as an Underlying Distribution	153
Sampling	154
Simple Random Sampling	155
Systematic Sampling	156
Cluster Sampling	157
Stratified Sampling	157
The Role of Sampling in Inferential Statistics	158
The Sampling Distribution of the Mean	158
Using the Normal Distribution as the Sampling Distribution of the Mean	160
Summary	162
Key Concepts	162
Exercises	163
Key Strokes for Selected Exercises	165

9

Hypothesis Testing: One-Sample Case for the Mean 169

Step 1: Stating the Hypothesis	170
Step 2: Computing the Test Statistic	172
Step 3: Setting the Criterion for Rejecting H_0	173
Errors in Hypothesis Testing	174
Level of Significance	175
Region of Rejection	176
Region of Rejection for Directional Alternate Hypotheses	179
Step 4: Deciding Whether to Reject or Retain the Null Hypothesis	182
Hypothesis Testing When σ^2 Is Unknown	184
Student's t Distribution	184
Degrees of Freedom	185
An Example Using the t Distribution	186
Statistical Significance versus Practical Importance	189
Summary	190
Key Concepts	191
Exercises	191
Key Strokes for Selected Exercises	193

10

Estimation: One-Sample Case for the Mean 197

Confidence Intervals 198

Level of Confidence 199

The Relationship Between Hypothesis Testing and Estimation 202

Statistical Precision in Estimation 203

Summary 204

Key Concepts 205

Exercises 205

Key Strokes for Selected Exercises 206

11

Hypothesis Testing and Estimation: Two-Sample Case for the Mean 209

Testing $H_0: \mu_1 - \mu_2 = 0$, Independent Samples 210

The Null Hypothesis 211

The Alternative Hypothesis 211

The Sampling Distribution of $(\bar{X}_1 - \bar{X}_2)$ for Independent Samples 211

A Computational Example 214

A Second Computational Example 215

Confidence Interval for $\mu_1 - \mu_2$ 216

Parametric Assumptions 217

Testing $H_0: \mu_1 - \mu_2 = 0$, Dependent Samples 217

A Computational Example 218

Confidence Interval for the Mean Difference with Dependent
Samples 220

Summary 221

Key Concepts 221

Exercises 222

Key Strokes for Selected Exercises 224

12

Hypothesis Testing for Other Statistics 227

Hypotheses About Proportions 228

Testing $H_0: P = a$ 228

Testing $H_0: P_1 - P_1 = 0$	230
Summary of Hypothesis Testing for Proportions	232
Hypotheses About Correlation Coefficients	233
Testing $H_0: \rho = a$	233
Testing $H_0: \rho = 0$	236
Testing $H_0: \rho_1 = \rho_2$ or $\rho_1 - \rho_2 = 0$	237
Summary of Hypothesis Testing for Correlation Coefficients	238
Hypotheses About Variances	239
Testing $H_0: \sigma^2 = a$	239
Testing $H_0: \sigma_1^2 = \sigma_2^2$	241
Summary of Hypothesis Testing for Variances	243
Summary	244
Key Concepts	244
Exercises	244
Key Strokes for Selected Exercises	247

13

Analysis of Variance 251

The Variables in ANOVA	252
Notation in ANOVA	253
Data Layout—An Example	253
The Logic of ANOVA	254
Partitioning the Sum of Squares	258
Two Variance Estimates	260
An ANOVA Example	261
Step 1: Stating the Hypothesis	261
Step 2: Computing the Test Statistic	261
Computational Formulas	261
Step 3: Setting the Criterion for Rejecting the Null Hypothesis	264
Step 4: Deciding Whether to Reject or Retain the Null Hypothesis	264
Assumptions Underlying ANOVA	265
Interpreting the Result—The Use of <i>Post Hoc</i> Multiple Comparison Tests	266
The Scheffé Method	266
The Tukey Method	268
ANOVA for Repeated Measures	269
Repeated-Measures ANOVA Example	271
Summary	273
Key Concepts	274
Exercises	274
Key Strokes for Selected Exercises	277