STATISTICS

An Introduction

Robert D. Mason
Douglas A. Lind
William G. Marchal

STATISTICS
An Introduction

Robert D. Mason
Douglas A. Lind
William G. Marchal
The University of Toledo

HARCOURT BRACE JOVANOVICH, PUBLISHERS
San Diego New York Chicago Atlanta Washington, D.C.
London Sydney Toronto

To Dorothy, Jane, and Andrea

Copyright © 1983 by Harcourt Brace Jovanovich, Inc.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Requests for permission to make copies of any part of the work should be mailed to: Permissions, Harcourt Brace Jovanovich, Inc., 757 Third Avenue, New York, N.Y. 10017.

ISBN: 0-15-583525-4
Library of Congress Catalog Card Number: 82-82707
Printed in the United States of America

Front endpapers: The Normal Probability Distribution table is from *CRC Standard Mathematical Tables,* Fifteenth Edition (West Palm Beach, Fla.: Copyright The Chemical Rubber Co., CRC Press, Inc.). Critical Values of Student's t Distribution table is from *The Ways and Means of Statistics* by Leonard J. Tashman and Kathleen R. Lamborn © 1979 by Harcourt Brace Jovanovich, Inc. Both are reproduced by permission.

Preface

This text has been designed for use in introductory statistics courses. Principles of both descriptive and inferential statistics are discussed, illustrated, and applied in situations close to most students' own experience. The mathematics can be handled easily by students with a limited background.

We have written and organized the material to provide the greatest possible flexibility of use in both content and format. The book's content is appropriate for any general statistics course. Illustrations and exercises are drawn from disciplines as varied as sociology, education, politics, demography, meteorology, sports, and mathematics. Considerable latitude has also been built into the organization of the chapter topics, so that the text can easily be adapted for use in one-semester, one-quarter, or two-quarter courses. A one-semester course, for example, might include Chapters 1–10, 13, and 14. Time permitting, Chapter 16 would be a logical addition.

A number of special features set this text apart in motivating and assisting students as they progress through the material:

- Each chapter begins with a set of performance objectives—what the reader should be able to do on completion of that chapter. The objectives serve the dual function of advance organizers and motivators.
- A brief chapter introduction reviews important concepts presented in the previous chapter and explains how they are linked to the present chapter. The introduction is followed by a chapter overview.
- The discussion of each concept is followed by a realistic statistical problem and its solution.

- A number of self-reviews are interspersed in each chapter. Each self-review is closely patterned after the chapter problems that preceded it. Answers and methods of solution are always shown in the margin. These self-reviews help students to monitor their own progress and provide them with constant reinforcement.
- Many interesting real-world exercises are incorporated in the body and at the end of each chapter. The answers and methods of solution for all even-numbered exercises are supplied in the book's appendix.
- Every important new term and formula is defined and placed in a box for easy reference.
- Both a summary and a chapter outline are included in every chapter. They are a valuable aid to students in pulling together a chapter's main ideas.
- An end-of-chapter achievement test covers all the material in that chapter, helping students to evaluate their own overall comprehension of the subject matter. Answers and methods of solution are provided in the appendix.
- A set of chapter highlights is included after each of six major groups of chapters. They include a brief review, key concepts, key terms, key symbols, review problems, and two ongoing cases (one dealing with a grocery store chain, and the other with a hospital) to be analyzed by the students.
- Throughout the book, a number of computer applications using the Statistical Package for the Social Sciences, MINITAB, and the BASIC programming language illustrate the computer's potential for problem solving.
- A complete glossary of terms and a set of standard statistical tables are also included in the book's appendix.
- The normal and t distributions are repeated inside the front cover for easy reference.
- In addition, a simple flowchart inside the back cover aids students in selecting the appropriate formula for computing the mean and the standard deviation.

A complete ancillary package accompanies the text. It includes a Study Guide, a Solutions Manual, and a set of suggested test questions for each chapter. The Study Guide is comprehensive, with an organization similar to that of the main text. Each chapter includes chapter objectives, an extensive summary, solved problems, student exercises with answers and methods of solution in the appen-

试读结束：需要全本请在线购买： www.ertongbook.com

dix, and tear-out chapter assignments to be graded by the instructor. Ample space is provided for computations. The Solutions Manual contains the full solutions to all exercises in the textbook and to all chapter assignments in the Study Guide.

Finally, we wish to acknowledge the valuable contribution made to this book by friends, students, colleagues, and collaborators throughout its various stages of development. In particular, we would like to express our thanks and appreciation to Richard J. Beres, John J. Bodner, Patricia S. DeJarnette, and Toni M. Somers at the University of Toledo, and Jerry Bergman, Spring Arbor College, for their helpful suggestions and assistance in class testing portions of the manuscript. Special thanks to Patricia E. Hetrick, who proofread the manuscript and checked the solutions, and to Dolores A. Lucitte and Jean K. Schaefer for typing some of the material.

We also wish to express our gratitude to John M. Rogers (California Polytechnic State University, San Luis Obispo), Bayard Baylis (The King's College, New York), William W. Lau (California State University, Fullerton), Kenneth R. Eberhard (Chabot College), and David Macky (San Diego State University), whose careful reviews made this a better text.

At Harcourt Brace Jovanovich, we are indebted to Terence W. Retchless, Susan Crosier, Jonathan Kroll, Sue Lasbury, Diane Polster, and Nancy A. Shehorn for their valuable assistance throughout the book's production. Special thanks to Liana Beckett for the outstanding job she did editing our manuscript.

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 Table IV from their book *Statistical Tables for Biological, Agricultural and Medical Research* (6th edition, 1974).

ROBERT D. MASON

DOUGLAS A. LIND

WILLIAM G. MARCHAL

Contents

Preface iii

1 **WHAT IS STATISTICS?** 1

Introduction 2
Types of Statistics 3
The Nature of Statistical Data 5
A Word of Encouragement 9

2 **SUMMARIZING DATA: FREQUENCY DISTRIBUTIONS AND GRAPHIC PRESENTATION** 11

Introduction 12
Frequency Distribution 12
Suggestions for Constructing a Frequency Distribution 17
Portraying the Frequency Distribution Graphically 19
Chapter Achievement Test 45

3 **DESCRIPTIVE STATISTICS: MEASURES OF CENTRAL TENDENCY** 49

Introduction 50
What Is an Average? 50

The Sample Mean 51
The Population Mean 53
The Properties of the Arithmetic Mean 54
The Median 57
The Mode 60
Choosing an Average 60
Computation of the Mean from Grouped Data 61
Computation of the Median from Grouped Data 67
The Mode from Grouped Data 71
Choosing an Average for Data in a Frequency Distribution 72
Chapter Achievement Test 81

4 DESCRIPTIVE STATISTICS: MEASURES OF DISPERSION AND SKEWNESS 85

Introduction 86
Measures of Dispersion 87
Measures of Dispersion for Data in a Frequency Distribution 98
Relative Dispersion 108
The Coefficient of Skewness 111
A Computer Example 112
Chapter Achievement Test 119
Highlights From Chapters 1, 2, 3, and 4 122

5 AN INTRODUCTION TO PROBABILITY 131

Introduction 132
Concepts in Probability 134
Types of Probability 136
Elementary Probability Rules 139
Some Principles of Counting 153
Chapter Achievement Test 166

6 PROBABILITY DISTRIBUTIONS 169

Introduction 170
What Is a Probability Distribution? 170
Discrete and Continuous Random Variables 173
The Binomial Probability Distribution 175
Chapter Achievement Test 186

7 THE NORMAL PROBABILITY DISTRIBUTION 189

Introduction 190
Characteristics of a Normal Probability Distribution 191
The "Family" of Normal Probability Distributions 192
Areas Under the Normal Curve 194
The Standard Normal Probability Distribution 195
The Normal Approximation to the Binomial 206
Chapter Achievement Test 218
Highlights From Chapters 5, 6, and 7 221

8 SAMPLING DISTRIBUTIONS AND THE CENTRAL LIMIT THEOREM 227

Introduction 228
Methods of Probability Sampling 229
Sampling Error 235
The Sampling Distribution of the Mean 236
The Central Limit Theorem 240
Confidence Intervals 241
Standard Error of the Mean 243
Point Estimates and Interval Estimates 244
Choosing an Appropriate Sample Size 247
Chapter Achievement Test 254

9 HYPOTHESIS TESTS: LARGE-SAMPLE METHODS 257

Introduction 258
The General Idea of Hypothesis Testing 259
A Test Involving the Population Mean (Large Samples) 262
A Test for Two Population Means (Large Samples) 271
Chapter Achievement Test 279

10 TESTING A HYPOTHESIS ABOUT PROPORTIONS 283

Introduction 284
A Test Dealing with Population Proportions 284
A Test Dealing with Two Population Proportions 289
Chapter Achievement Test 297
Highlights From Chapters 8, 9, and 10 300

11 HYPOTHESIS TESTS: SMALL-SAMPLE METHODS 305

Introduction 306
Characteristics of the t Distribution 306
Testing a Hypothesis About a Population Mean 308
Comparing Two Population Means 315
A Test of Paired Observations 321
Chapter Achievement Test 330

12 ANALYSIS OF VARIANCE 335

Introduction 336
Underlying Assumptions for ANOVA 336

The ANOVA Test 339
Chapter Achievement Test 356
Highlights From Chapters 11 and 12 **359**

13 CORRELATION ANALYSIS 367

Introduction 368
Correlation Analysis 368
The Coefficient of Correlation 373
The Coefficient of Determination 378
Testing the Significance of the Coefficient of Correlation 380
Rank-Order Correlation 384
Chapter Achievement Test 395

14 REGRESSION ANALYSIS 399

Introduction 400
Linear Regression 400
Fitting a Line to Data 406
The Method of Least Squares 410
The Standard Error of Estimate 416
Assumptions About Regression 420
Computing Confidence-Interval Estimates 421
Chapter Achievement Test 438

15 MULTIPLE REGRESSION AND CORRELATION ANALYSIS 441

Introduction 442
Assumptions About Multiple Regression and Correlation 442
The Multiple Regression Equation 443
The Multiple Standard Error of Estimate 447
Coefficients of Multiple Correlation and Determination 449

试读结束：需要全本请在线购买： www.ertongbook.com

A Computer Application Using the Stepwise Multiple Regression Procedure 449
Chapter Achievement Test 461
Highlights From Chapters 13, 14, and 15 465

16 ANALYSIS OF NOMINAL-LEVEL DATA: THE CHI-SQUARE DISTRIBUTION 473

Introduction 474
The Chi-Square Distribution 474
Goodness-of-Fit Test: Equal Expected Frequencies 476
Goodness-of-Fit Test: Unequal Expected Frequencies 481
Contingency-Table Analysis 487
Chapter Achievement Test 498

17 NONPARAMETRIC METHODS: ANALYSIS OF RANKED DATA 501

Introduction 502
The Sign Test 502
The Wilcoxon Signed-Rank Test 513
The Wilcoxon Rank-Sum Test 518
Analysis of Variance by Ranks (the Kruskal-Wallis Test) 523
Chapter Achievement Test 531
Highlights From Chapters 16 and 17 534

ANSWERS TO CHAPTER ACHIEVEMENT TESTS 541

SOLUTIONS TO EVEN-NUMBERED EXERCISES 553

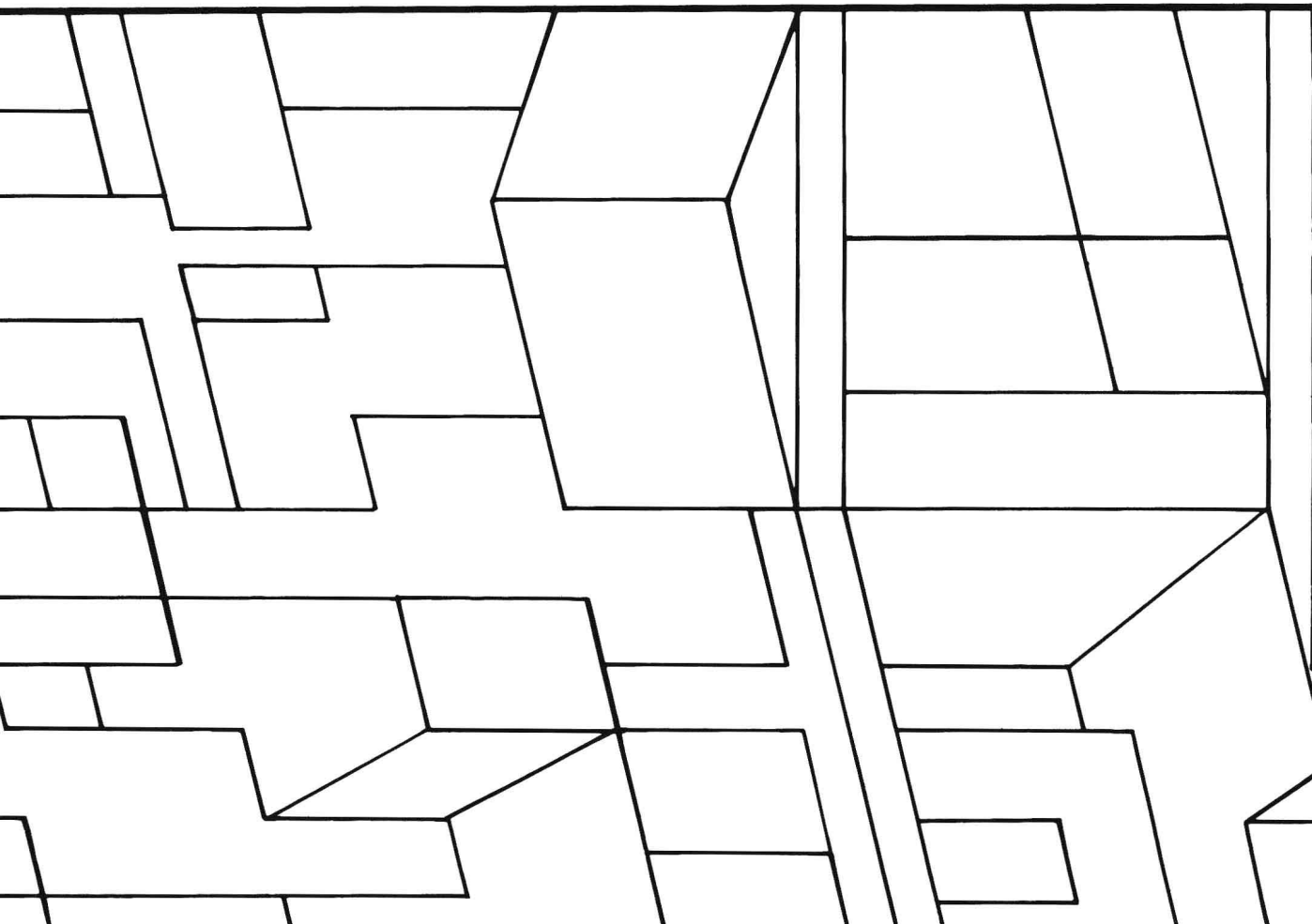
APPENDIX TABLES 601

A The Binomial Probability Distribution 601
B A Table of Random Numbers 610
C The Normal Probability Distribution 612
D Critical Values of Student's t Distribution 613
E Critical Values of the F Statistic 614
F Critical Values of the Chi-Square Statistic 616
G Critical Values of the Wilcoxon Signed-Rank Statistic 617

GLOSSARY 618

INDEX 623

1
What Is Statistics?



Introduction

Whenever we watch television, listen to the radio, or read a newspaper or magazine, all of us are exposed to—and sometimes overwhelmed by—assorted facts and figures commonly labelled "statistics." We may read or hear that

- The Consumer Price Index increased 0.8% last month.
- More than 97% of all households in the United States have television sets.
- One violent crime is committed every 27 seconds.
- The Dallas Cowboys defeated the Los Angeles Rams 29 to 17. The Cowboys had 24 first downs, the Rams 22. The passing yards were 275 for Dallas, 203 for Los Angeles. Dallas was penalized 108 yards, Los Angeles 78 yards.

As a result, you may envision a statistician as someone who sits in the press box and collects numerical data on passing yardage, punt returns, and so on and releases statistics like these:

Rams - Cowboys

October 18

Score by Periods

Los Angeles . .	0	10	7	0—17
Dallas	12	14	0	3—29

Scoring

Dallas—Springs 1 run (Septien kick).
Dallas—Field goal Septien 40.
Dallas—Safety, Haden tackled in end zone.
Dallas—Dorsett 44 run (Septien kick).
Los Angeles—Tyler 2 run (Corral kick).
Los Angeles—Field goal Corral 40.
Dallas—Hill 63 pass from D. White (Septien kick).
Los Angeles—D. Hill 43 pass from Haden (Corral kick).
Dallas—Field goal Septien 39.

Team Statistics

	Los Angeles	Dallas
First downs . .	22	24
Rushes - Yards . . .	34 - 171	42 - 221

Individual Statistics

Rushing—Dallas, Dorsett 27 - 159, Springs 12 - 41, D. White 3 - 21; Los Angeles, Tyler 16 - 90, Bryant 14 - 53, Haden 2 - 17, Guman 2 - 11.

Passing—Dallas, D. White 15 - 33 - 2 —277; Los Angeles, Haden 13 - 30 - 3 —237.

Receiving—Dallas, Hill 4 - 97, Pearson 4 - 78, Saldi 2 - 17, Cosbie 1 - 28, Dorsett 1 - 22, DuPree 1 - 15, Johnson 1 - 15, Springs 1 - 5; Los Angeles, Miller 4 - 51, Moore 3 - 53, D. Hill 2 - 88, Bryant 2 - 13, Childs 1 - 19, Arnold 1 - 13.

Kickoff Returns—Los Angeles, D. Hill 6 - 95; Dallas, J. Jones 2 - 40, Fellows 1 - 14.

Punt Returns—Los Angeles, Irvin 3 - 13; Dallas, J. Jones 1 - 9, Fellows 2 - 6.

Interceptions—Los Angeles, Cromwell 2 - 0; Dallas, Walls 2 - 0, Thurman 1 - 0.

Punting—Los Angeles, Corral 5 - 48.0; Dallas, D. White 5 - 42.6.

Such sports data *are,* in fact, one variety of statistics—a collection of numerical data. However, statistics also has a broader meaning. The boxed definition that follows is the one we will consider throughout.

> **Statistics** The body of techniques used to facilitate the collection, organization, presentation, analysis, and interpretation of data for the purpose of making better decisions.

Types of Statistics

Descriptive Statistics

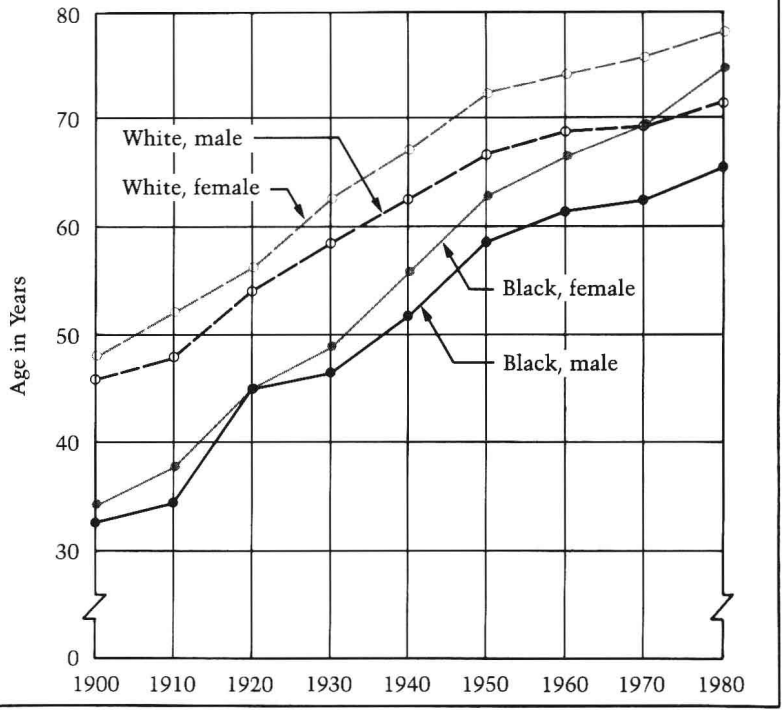
As stated in the definition, statistics involves, among other things, the collection, organization, and presentation of numerical data. Masses of unorganized data stored in a computer, or collected by Gallup-type polls with respect to the preference of voters, are usually of little value. Techniques are available, however, to organize such data into some meaningful form. These aids in organizing, analyzing, and describing (or summarizing) a large collection of numbers are collectively referred to as **descriptive statistics.**

> **Descriptive Statistics** Methods used to describe the data that have been collected.

We will discuss one such summary technique, called a frequency distribution, in Chapter 2. In that chapter, we will also examine how data can be presented graphically. For example, over 300 separate figures were needed to create a chart of a population's life expectancy at birth, by ethnic background and sex, since 1900 (see Figure 1-1). At a single glance, this chart reveals that the life expectancy for all groups has risen dramatically since 1900—with the life expectancy for blacks having increased over 100%.

Figure 1-1

Life Expectancy at Birth, by Ethnic Background and Sex: U.S. 1900-1980



In Chapter 3 we will study the tendency of numerical data to cluster about a central value. An analysis of the spread or variation in the data using measures such as the standard deviation and the variance will be found in Chapter 4.

Inferential Statistics

Descriptive statistics is only one facet of the science of statistics. Another is **inferential statistics** or **inductive statistics.** Inferential statistical methods are concerned with finding out something about a **population.** Gallup, Harris, and other pollsters do just that when they are hired before an election to find out how voters (the population) plan to vote on election day.

试读结束：需要全本请在线购买： www.ertongbook.com