

MENDENHALL

*Seventh Edition*

INTRODUCTION TO  
PROBABILITY AND STATISTICS



SEVENTH  
EDITION

# INTRODUCTION TO PROBABILITY AND STATISTICS



*William Mendenhall*



DUXBURY PRESS  
*Boston*

# PWS PUBLISHERS

Prindle, Weber & Schmidt • Duxbury Press • PWS Engineering • Breton Publishers  
20 Park Plaza • Boston, Massachusetts 02116

Copyright © 1987, 1983 by PWS Publishers.

Copyright © 1979 by Wadsworth Publishing Company, Inc.

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transcribed, in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written permission of PWS Publishers.

PWS Publishers is a division of Wadsworth, Inc.

## *Library of Congress Cataloging-in-Publication Data*

Mendenhall, William  
Introduction to probability and statistics.

Includes bibliographies and index.

1. Mathematical statistics. 2. Probabilities.

I. Title.

QA276.M425 1987 519.2 86-19690

ISBN 0-87150-046-9

## *Credits*

Exercises 8.15 (p. 314), 10.113 (p. 470), 10.115 (p. 471), 11.14 (p. 495), and 11.64 (p. 521) reprinted by permission of the American Alliance for Health, Physical Education, Recreation and Dance, 1900 Association Drive, Reston, VA 22091.

Exercise 9.21 (p. 370) reprinted with permission from *Psychology Today* magazine. Copyright © 1981 by the American Psychological Association.

Exercises 10.59 (p. 434) and 10.150 (pp. 478–479) reprinted by permission of Elsevier Science Publishing Co., Inc. Copyright 1983 by The American Psychosomatic Society, Inc.

Exercises 9.56 (p. 388) and 11.50 (p. 513) copyright © 1981 by the American Institute of Biological Sciences. Reprinted with permission of the copyright holder.

Printed in the United States of America.

87 88 89 90 91–10 9 8 7 6 5 4 3 2 1

**Sponsoring Editor:** Michael Payne

**Production Coordinator/Designer:** Elise Kaiser

**Production:** Del Mar Associates

**Composition:** Weimer Typesetting Company, Inc.

**Cover Printer:** New England Book Components

**Text Printer/Binder:** R. R. Donnelley & Sons Company

Cover art "Autumn Garden with Moat" 1981 by Taj Worley; a lithograph on handmade paper, used with the permission of the artist.

# PREFACE



The teaching objective of this seventh edition of *Introduction to Probability and Statistics* is to provide a connected introduction to statistics that presents statistical inference, *the objective of statistics*, as its theme. To achieve this objective, the seventh edition retains the features of earlier editions, slightly modified, with some substantial changes. These changes include the following.

## *Changes in This Edition*

1. "Useful Mathematical Notation" (formerly Chapter 2) has been moved to the back of the text as Appendix II.
2. A data set consisting of the diastolic and systolic blood pressures of 966 males and 945 females, ages 15 through 20, has been added as Appendix I. These data, which represent part of a large study conducted by the National Institutes of Health, form the basis of the case study in Chapter 2, and they are used in exercises throughout the text.
3. Stem and leaf diagrams and the box plot method for identifying outliers have been added to Chapter 2.
4. Old Chapter 7, which included material on the normal distribution and sampling distributions, has been split into two chapters: Chapter 6 (the normal distribution) and Chapter 7 (sampling distributions). The material on sampling distributions has been substantially increased.
5. The Poisson approximation to the binomial probability distribution has been added as an optional section in new Chapter 6.
6. Chapter 12, a multiple regression analysis, has been improved by the addition of three new examples early in the chapter. Data are presented implying a curvilinear relationship between  $y$  and  $x$ . Second-order and exponential models are proposed for fitting the

### *Features of This Text*

data, and the computer outputs of the regression analyses are discussed for each. A brief section on a stepwise regression analysis has also been added, as well as a section identifying the problems introduced by model selection, multicollinearity, and so forth.

7. The Kruskal-Wallis  $H$  test and the Friedman  $F_r$  test, nonparametric tests used for detecting differences in locations of more than two population distributions, have been added to Chapter 16. These tests provide nonparametric alternatives to the parametric analyses of variance for the completely randomized and randomized block designs of Chapter 15.

8. Three new case studies, which introduce the chapters and are then discussed in separate sections, have been added.

9. Many new applied exercises, based on newspaper, magazine, or research articles, have been added to provide a varied assortment of 1,300 exercises from which to choose. Exercise sets at the ends of sections have been divided into those that deal with straightforward applications of the methodology (basic techniques) and those that are "word problems" and are concerned with applications.

1. **Connectivity:** The goal of the text is to present an introduction to statistical inference. Each chapter begins by explaining how the chapter material plays a role in achieving this objective and why its study is necessary. At various points in the text we take time to tie the chapter topics together and to summarize how they play a role in statistical inference.

2. **Motivation:** The author believes that there are two ingredients needed to motivate students to study statistics. They need to be able to understand how statistics works and to see how it can be applied in the real world of work.

The connective approach to the subject of statistical inference, repetition of the objective of statistics, and repeated explanations of how various topics support this objective help to provide understanding of the subject. The varied applications of statistical inference, in the biological, social, and physical sciences, in business, engineering, medicine, and law, provide convincing evidence that statistics is used in the real world and that learning statistics is worthwhile.

3. **Pedagogy:** Some of the pedagogical devices employed in this text have just been described. In addition to these, case studies—human interest applications of statistics that introduce each chapter—are provided to pose a question and to stimulate interest in the chapter's topic. After the student has covered the appropriate chapter material, we show (in a separate section) how the chapter methodology can be used to answer the case study question.

## *Exercises and Other Aids*

## *Acknowledg- ments*

Definitions and methods are boxed. Key sentences in the body of the text (the type outlined by some students with a highlighter) are shown in bold with a color bar to the left to assist in review and study. Tips on problem solving are boxed and inserted at key points to improve the student's problem-solving techniques.

Exercises are graduated in difficulty so that all students can solve some of them, a substantial number can solve most, and a few of the best students are challenged to solve all without error. Symbols are used to identify areas of application in business, industry, and the sciences. A thorough knowledge of definitions and concepts is essential before students attempt the exercises. Unannounced quizzes are very effective in helping the student learn the new language of statistics.

A semiprogrammed study guide to help students who have individual difficulties with the subject matter is available. This study guide, now in its seventh edition, has benefited from considerable user experience. An exercise solutions manual for instructors and a partial solutions manual for students are also available.

The preparation of a text and its subsequent new editions is always a team effort, and I am fortunate to have had the assistance of many talented people. The late William Miller and Paul Benson of Bucknell University helped develop and test the first edition and made comments concerning its revision. Numerous reviewers aided in the writing of the second, third, fourth, fifth, and sixth editions. They include John T. Webster, William Brelsford, Robert Crovelli, Douglas Chapman, Arthur Coladarci, the late Paul Meyer, Joyce Curry, Frank Deane, Roy E. Myers, P. V. Rao, Dennis D. Wackerly, and John N. Quiring. The following reviewers provided detailed comments on this manuscript: Michael Burke, College of San Mateo; S. Cheng, University of Manitoba; Dale O. Everson, University of Idaho; Rudy J. Freund, Texas A & M; David Groggel, Miami University; Donald K. Hotchkiss, Iowa State University; Kermit Hutcheson, University of Georgia; Donald R. LaTorre, Clemson University; Donald Ramirez, University of Virginia; Kirk Steinhorst, University of Idaho; Bill Stines, North Carolina State University; Leonard Sweet, The University of Akron; and Shie-Shen Yang, Kansas State University. A special note of appreciation is extended to Barbara Beaver for detailed line-by-line critique and for preparing the solutions manual for this and other editions. I would also like to express my appreciation to the production and editorial staffs of Duxbury Press. Thanks are also due to E. J. Pearson, A. Hald, W. H. Beyer, R. A. Wilcox, and the Biometrika Trustees for their kind permission to use tables reprinted in Appendix III. I am indebted to the typists who have given their best and endured the worst: Florence Valen-

tine, who typed the original draft of the manuscript; Angie Anastasia and Gay Midelis, who prepared the second revision; Mary Jackson and Catherine Kennedy, who typed the third edition; Ellen Evans, who typed the fourth; Carol Rozear, who typed the fifth edition; Brenda Dobson, who typed the sixth edition; and Carol Springer, who typed drafts of the current edition. Finally, as before, I acknowledge the assistance of my family and their partnership in this writing endeavor.

*William Mendenhall*

# CONTENTS



## CHAPTER 1

### *What Is Statistics?* 1

#### *CASE STUDY Reaping the Rewards of Refunds* 2

- 1.1 Illustrative Statistical Problems 3
- 1.2 The Population and the Sample 4
- 1.3 The Essential Elements of a Statistical Problem 5
- 1.4 The Role of Statistics and the Statistician 7
- 1.5 More on the Consumer Refund Case Study 7
- 1.6 Summary 8
- 1.7 A Note to the Reader 9

## CHAPTER 2

### *Describing Sets of Data* 14

#### *CASE STUDY How Is Your Blood Pressure?* 15

- 2.1 A Graphical Method for Describing a Set of Data: Relative Frequency Distributions 16
- 2.2 Stem and Leaf Displays (Optional) 26
- 2.3 Numerical Methods for Describing a Set of Data 30
- 2.4 Measures of Central Tendency 31
- 2.5 Measures of Variability 33
- 2.6 On the Practical Significance of the Standard Deviation 38
- 2.7 A Short Method for Calculating the Variance 46
- 2.8 A Check on the Calculation of  $s$  49



## CHAPTER 3

- 2.9 Measures of Relative Standing 56
- 2.10 Blood Pressure: How You Compare with Your Contemporaries 60
- 2.11 Outliers: z-Scores and Box Plots (Optional) 63
- 2.12 Summary and Comments 68

*Probability* 74*CASE STUDY Probability and Decision Making in the Congo* 75

- 3.1 The Role of Probability in Statistics 76
- 3.2 Two Teaching Objectives 77
- 3.3 The Sample Space 78
- 3.4 The Probability of an Event 81
- 3.5 Compound Events 89
- 3.6 Event Relations 93
- 3.7 Two Probability Laws and Their Use 96
- 3.8 More on the ERTS Case Study 110
- 3.9 Bayes' Rule (Optional) 112
- 3.10 Results Useful in Counting Sample Points (Optional) 116
- 3.11 Summary 129

## CHAPTER 4

*Random Variables and Probability Distributions* 137*CASE STUDY Thyroid Disease and Three Mile Island* 138

- 4.1 Random Variables 139
- 4.2 A Classification of Random Variables 140
- 4.3 Probability Distributions for Discrete Random Variables 142
- 4.4 Probability Distributions for Continuous Random Variables 149
- 4.5 Mathematical Expectation 151
- 4.6 Analysis of the Number of Hyperthyroid Births near Three Mile Island 163
- 4.7 Random Sampling 164
- 4.8 Summary 167

## CHAPTER 5

*The Binomial Probability Distribution* 173*CASE STUDY The Fear of Flying* 174

- 5.1 The Binomial Experiment 175
- 5.2 The Binomial Probability Distribution 177
- 5.3 The Mean and Variance for the Binomial  
Random Variable 188
- 5.4 More on the Fear of Flying Survey 194
- 5.5 Making Decisions: Lot Acceptance Sampling for  
Defectives 195
- 5.6 Making Decisions: A Test of an Hypothesis 201
- 5.7 Choosing the Null Hypothesis 206
- 5.8 A General Comment 207
- 5.9 Summary 214

## CHAPTER 6

*The Normal Probability Distribution* 219*CASE STUDY The Long and the Short of It* 220

- 6.1 The Normal Probability Distribution 221
- 6.2 Tabulated Areas of the Normal Probability  
Distribution 222
- 6.3 Fortune's Forty-to-One Odds 232
- 6.4 The Normal Approximation to the Binomial  
Distribution 234
- 6.5 The Poisson Approximation to the Binomial Probability  
Distribution (Optional) 246
- 6.6 Summary 252

## CHAPTER 7

*Sampling Distributions* 258*CASE STUDY Sampling the Roulette at Monte Carlo* 259

- 7.1 Sampling Distributions of Statistics 260
- 7.2 The Central Limit Theorem and the Sampling Distribution of  
the Sample Mean 261
- 7.3 The Sampling Distribution of a Sample Proportion 276
- 7.4 The Sampling Distribution of the Sum or of the Difference  
Between Two Independent Statistics 282

## CHAPTER 8

- 7.5 The Sampling Distribution of Winnings at Roulette 291
- 7.6 Summary 292

*Large-Sample Estimation* 296*CASE STUDY How Reliable Is That Poll?* 297

- 8.1 A Brief Summary 298
- 8.2 Inference: The Objective of Statistics 298
- 8.3 Types of Estimators 300
- 8.4 Point Estimation of a Population Mean 303
- 8.5 Interval Estimation of a Population Mean 307
- 8.6 Estimation from Large Samples 316
- 8.7 Estimating the Difference Between Two Means 317
- 8.8 Estimating the Parameter of a Binomial Population 322
- 8.9 Estimating the Difference Between Two Binomial Parameters 327
- 8.10 Choosing the Sample Size 331
- 8.11 How Reliable Is That *New York Times*/WCBS-TV Opinion Poll? 336
- 8.12 Summary 340

## CHAPTER 9

*Large-Sample Tests of Hypotheses* 347*CASE STUDY Measuring the Effects of Passive Smoking* 348

- 9.1 Testing Hypotheses about Population Parameters 349
- 9.2 A Statistical Test of an Hypothesis 350
- 9.3 A Large-Sample Statistical Test 353
- 9.4 Testing an Hypothesis about a Population Mean 356
- 9.5 Testing an Hypothesis about the Difference Between Two Population Means 362
- 9.6 Analyzing the Passive Smoking Data 363
- 9.7 Testing an Hypothesis about a Population Proportion 372
- 9.8 Testing an Hypothesis about the Difference Between Two Population Proportions 378
- 9.9 Another Way to Report the Results of the Statistical Tests: *p*-Values 384

## CHAPTER 10

9.10 Some Comments on the Theory of Tests of Hypotheses 389

9.11 Summary 390

*Inference from Small Samples* 397

*CASE STUDY Just a Pinch Between the Cheek and the Gum* 398

10.1 Introduction 399

10.2 Student's  $t$  Distribution 399

10.3 Small-Sample Inferences Concerning a Population Mean 403

~~10.4~~ Small-Sample Inferences Concerning the Difference Between Two Means 412

10.5 A Paired-Difference Test 425

10.6 More on Smokeless Tobacco 435

10.7 Inferences Concerning a Population Variance 441

10.8 Comparing Two Population Variances 450

10.9 Assumptions 461

10.10 Summary 463

## CHAPTER 11

*Linear Regression and Correlation* 481

*CASE STUDY Does It Pay to Save?* 482

11.1 Introduction 484

11.2 A Simple Linear Probabilistic Model 485

11.3 The Method of Least Squares 489

11.4 Calculating  $s^2$ , an Estimator of  $\sigma^2$  496

11.5 Inferences Concerning the Slope of the Line,  $\beta_1$  499

11.6 Estimating the Expected Value of  $y$  for a Given Value of  $x$  509

11.7 Predicting a Particular Value of  $y$  for a Given Value of  $x$  515

11.8 A Coefficient of Correlation 521

11.9 A Multivariable Predictor 530

11.10 Case Study: A Computer Analysis of Whooping Crane Sightings 531



## CHAPTER 12

- 11.11 Assumptions 536
- 11.12 Does the Personal Savings Rate Increase as Investment  
Income Tax Liability Increases? 537
- 11.13 Summary 540

*Multiple Regression Analysis* 550*CASE STUDY Predicting Worker Absenteeism* 551

- 12.1 The Objective of a Multiple Regression Analysis 552
- 12.2 The Multiple Regression Model and Associated  
Assumptions 553
- 12.3 A Multiple Regression Analysis 556
- 12.4 A Comparison of Computer Printouts 570
- 12.5 Stepwise Regression Analysis 577
- 12.6 Misinterpretations in Regression Analyses 579
- 12.7 Some Comments on Model Formulation (Optional) 580
- 12.8 A Multiple Regression Analysis for the Worker Absenteeism  
Case Study 583
- 12.9 Summary 587

## CHAPTER 13

*Analysis of Enumerative Data* 597*CASE STUDY No Wine Before Its Time* 598

- 13.1 A Description of the Experiment 599
- 13.2 The Chi-Square Test 600
- 13.3 A Test of an Hypothesis Concerning Specified Cell  
Probabilities 602
- 13.4 Contingency Tables 607
- 13.5  $r \times c$  Tables with Fixed Row or Column Totals 617
- 13.6 A Computer Analysis for an  $r \times c$  Contingency Table 622
- 13.7 Wine Preferences: Canned or Bottled? 624
- 13.8 Other Applications 626
- 13.9 Assumptions 628
- 13.10 Summary 629

## CHAPTER 14

### *Considerations in Designing Experiments* 638

#### *CASE STUDY Does Depression Cause You to Eat More?* 639

- 14.1 The Factors Affecting the Information in a Sample 640
- 14.2 The Physical Process of Designing an Experiment 641
- 14.3 Random Sampling and the Completely Randomized Design 644
- 14.4 Volume-Increasing Experimental Designs 646
- 14.5 Noise-Reducing Experimental Designs 651
- 14.6 An Example of a Factorial Experiment 655
- 14.7 Summary 658

## CHAPTER 15

### *The Analysis of Variance* 661

#### *CASE STUDY A Comparison of Car-Insurance Costs for Different Locales* 662

- 15.1 Introduction 663
- 15.2 The Analysis of Variance 663
- 15.3 A Comparison of More Than Two Means 671
- 15.4 Proof of Additivity of the Sums of Squares (Optional) 675
- 15.5 An Analysis-of-Variance Table for a Completely Randomized Design 676
- 15.6 Estimation for the Completely Randomized Design 677
- 15.7 A Computer Printout for a Completely Randomized Design 679
- 15.8 The Analysis of Variance for a Randomized Block Design 688
- 15.9 Estimation for the Randomized Block Design 692
- 15.10 A Computer Printout for a Randomized Block Design 694
- 15.11 An Analysis of the Difference in Car-Insurance Costs for Four Locales 701
- 15.12 The Analysis of Variance for a Latin Square Design 702
- 15.13 Estimation for the Latin Square Design 707
- 15.14 A Computer Printout for a Latin Square Design 708
- 15.15 Selecting the Sample Size 713

## CHAPTER 16

- 15.16 Some Cautionary Comments on Blocking 717
- 15.17 Assumptions 717
- 15.18 Summary and Comments 719

## *Nonparametric Statistics* 727

### *CASE STUDY Does It Pay to Save? A Second Look* 728

- 16.1 Introduction 729
- 16.2 The Sign Test for Comparing Two Populations 730
- 16.3 A Comparison of Statistical Tests 736
- 16.4 The Mann-Whitney  $U$  Test: Independent Random Samples 736
- 16.5 The Wilcoxon Signed-Rank Test for a Paired Experiment 747
- 16.6 The Kruskal-Wallis  $H$  Test for Completely Randomized Designs 754
- 16.7 The Friedman  $F_r$  Test for Randomized Block Designs 760
- 16.8 Rank Correlation Coefficient 764
- 16.9 Does It Pay to Save? A Rank Correlation Analysis 773
- 16.10 Summary 775

## CHAPTER 17

### APPENDIX I

### ★ APPENDIX II

### APPENDIX III

## *A Summary and Conclusion* 781

### *NIH Blood Pressure Data Set* A-1

### *Useful Mathematical Notation* A-16

### *Tables* A-29

### *Answers* A-60

### *Index* A-93

## CHAPTER

# I

### *General Objective*

### *Specific Objectives*

## WHAT IS STATISTICS?



The purpose of this chapter is to identify the nature of statistics, its objective, and how it plays an important role in the sciences, in industry, and, ultimately, in our daily lives.

1. To use a case study to motivate a discussion of statistics.
2. To answer the question, What is statistics? *Sections 1.1, 1.2, 1.3, 1.4*
3. To identify statistical inference as the objective of modern statistics. *Section 1.3*
4. To identify the contributions statistics can make to inference making based on sample data. *Section 1.4*
5. To apply the knowledge gained in this chapter to the case study. *Section 1.5*
6. To define the basic words and concepts used in statistics. *Sections 1.1–1.7*



## CASE STUDY

# REAPING THE REWARDS OF REFUNDS

Will court-enforced retail refunds make you rich? Not likely, according to a *Wall Street Journal* article (June 9, 1981) that reports on a court case stemming from a price-fixing indictment. According to the article, approximately 54,000 persons—customers of Saks Fifth Avenue, Bonwit Teller, and Bergdorf Goodman during the period 1968 to 1974—may be due refunds if they can resurrect their receipts and file the necessary applications. The case raises a serious question about this and other court-directed refund programs that are often the result of price-fixing cases. Who profits (other than the attorneys involved in the lawsuit)? Specifically, how many of the 54,000 customers eligible for refunds will actually find their receipts and submit the required applications? And how much of the settlement money will really be distributed to the eligible customers? To answer these questions, you would need to either interview all 54,000 customers who are potentially eligible (an impossible task) or select a sample from the 54,000 and use statistical methods to estimate the number of people eligible for refunds and the amount of money that will eventually be refunded.

Our intention here is not to answer the question, Who profits from refunding? but to present a difficult problem that can be solved using a very powerful tool, statistics. In this chapter we will describe the objective of statistics. In particular we will identify the types of problems that statistical methodology can solve and explain how this valuable tool can be used to answer some very practical questions. We will return to the refunding case study in Section 1.5.