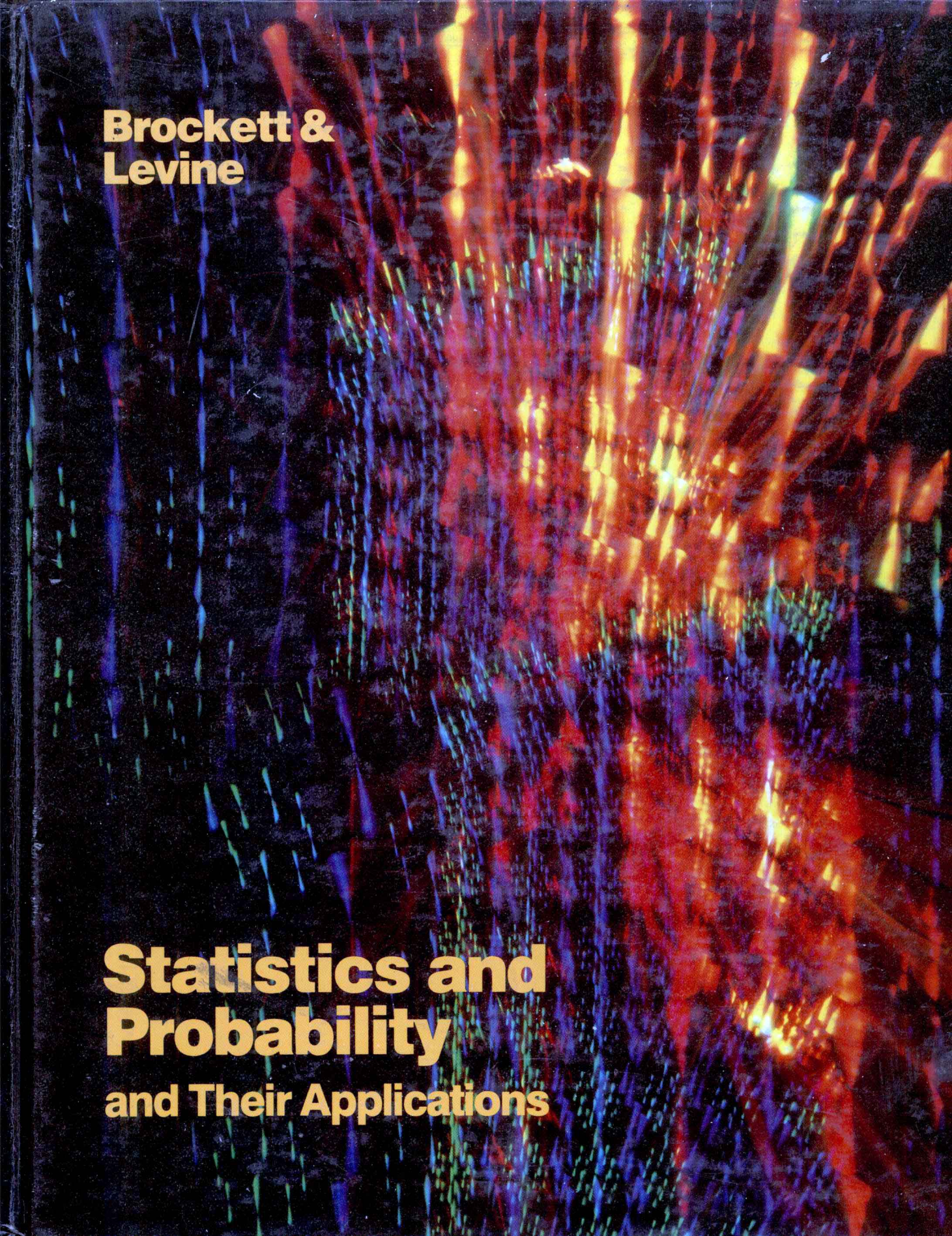
Brockett &
Levine
Statistics and
Probability
and Their Applications

STATISTICS & PROBABILITY & THEIR APPLICATIONS

Patrick Brockett
University of Texas at Austin

Arnold Levine
Tulane University

SAUNDERS COLLEGE PUBLISHING

Philadelphia New York Chicago
San Francisco Montreal Toronto
London Sydney Tokyo Mexico City
Rio de Janeiro Madrid

Address orders to:
383 Madison Avenue
New York, NY 10017

Address editorial correspondence to.
West Washington Square
Philadelphia, PA 19105

Text Typeface: 10/12 Times Roman
Compositor: Progressive Typographers
Acquisitions Editor: Leslie Hawke
Developmental Editor: Lloyd Black
Project Editor: Robin C. Bonner
Copyeditor: Elaine L. Honig
Art Director: Carol C. Bleistine
Art/Design Assistant: Virginia A. Bollard
Text Design: Lisa Delgado
Cover Design: Lawrence R. Didona
Text Artwork: Philatek
Managing Editor & Production Manager: Tim Frelick
Assistant Production Manager: Maureen Iannuzzi

Cover credit: Abstract. copyright © G. Gove/The Image Bank

Library of Congress Cataloging in Publication Data

Brockett, Patrick.
Statistics and probability and their applications.

Includes index.
1. Mathematical statistics. 2. Probabilities.
I. Levine, Arnold. II. Title.
QA276.B69 1984 519.5 84-5384
ISBN-0-03-053406-2

STATISTICS AND PROBABILITY AND THEIR APPLICATIONS ISBN 0-03-053406-2

© 1984 by CBS College Publishing.

All rights reserved. Printed in the United States of America.
Library of Congress catalog card number 84-5384

567 016 98765432

CBS COLLEGE PUBLISHING
Saunders College Publishing
Holt, Rinehart and Winston
The Dryden Press

Preface

A noncalculus statistics text of the kind that is useful to both students and practitioners in a wide variety of fields faces a number of stringent demands. The traditional material (random variables, estimation, hypothesis testing, regression, correlation, and so on) must be covered together with the appropriate notation so that a student who wishes to continue along either the mathematical road or any other route to learning statistics is in no way penalized. Moreover, the material must be written to take into account the considerable differences between readers with respect to their facility with basic algebra.

In addition, two other demands are increasingly being made by those departments that desire their students to obtain a basic foundation in statistics. First, they insist that their students *understand* the material. The applications of statistics studied by undergraduates in their special fields have become sufficiently profound that a "cookbook" course no longer suffices as background. The other demand concerns the *coverage* of the material. Generally, textbooks emphasize the analysis of data. The design of the data collection process, the checking, editing, and reduction of data generally merit little attention. By not being exposed to these methodologies, the student obtains a distorted view of statistical activities. As a result, there is an increasing demand that nonsuperficial discussions of sampling techniques and sampling instruments (such as questionnaires) be included in a first course.

We have attempted to satisfy the conditions noted above. A careful attempt has been made to explain notation and conventions clearly; we have included discussions of "data collection" techniques, including questionnaire design, survey and sampling design, experimental design, and so on, to provide a balance of coverage not available in texts of this level. The area of "data collection" is a large one; we have attempted to give the student insight into those techniques used both in scientific experiments and in surveys.

In addition, we have taken pains to include a great deal of material of importance to practitioners. We present examples and problems taken from actual situations in agronomy, business, biology, economics and finance, education,

engineering, epidemiology, law, medical research, social science, physics, and psychology. The inclusion of numerous applications should give the reader a strong sense of "applied statistics" and through this sense a better understanding of statistical theory.

We have also sought to increase this understanding by using the concept of the "mathematical model" to present efficiently the basic theory. This concept facilitates both the discussion of probability and the models underlying the statistical analyses.

The text has been designed to give the instructor considerable latitude in the organization of course material. Some independent pathways are plotted in the course outlines that follow. One can begin, for example, with probability, the preference of many instructors, rather than with the "presentation of data." The instructor, the one most knowledgeable about the strengths and weaknesses of the students, may find none of these particular suggestions applicable. There is sufficient material to derive many other course sequences.

ORGANIZATION OF COURSES

The following are suggestions for course outlines. However, the instructor may have a time schedule or a need for emphasis totally different from that described below. For example, one of the authors had to present a two-week course for quality-control engineers at an aircraft plant in Los Angeles; this course had to be integrated with other two-week courses in physics and engineering given by different instructors. In this case, none of the outlines below were directly applicable, yet the material selected for this short statistics course was gleaned from the material in this text.

I. One-Year Course (2 semesters)

   A. Emphasis on statistics, with light coverage of probability.
      1. First semester: Chapters 1 to 9. Omit sections 3.8, 4.8 to 4.11, 4.14 to 4.16. Assign the easiest problems in Chapters 3, 4, and 5. If 4.15 is omitted, then sections 7.7 and 8.6 should be omitted also.
      2. Second semester: Chapters 10 to 15. Note that Chapters 8, 10, and 11 have sections that discuss data collection in scientific experiments whereas Chapters 13 and 14 deal mostly with survey sampling. The instructor can select those topics that are of interest to the type of student taking the course since these topics are discussed independently of each other.

   B. Emphasis on probability.
      1. First semester: Chapters 3 to 7. Chapter 2 may be included if desired. Explain all models including those in the problem sets of Chapters 3 and 4.

此为试读,需要完整PDF请访问: www.ertongbook.com

2. Second semester: Chapters 7 to 13. Sections 13.3 to 13.6 may be omitted if the instructor wishes to have time to include sections of interest from Chapters 14 and 15.

C. A course intermediate between A and B.
   1. First semester: Chapters 3 to 8 omitting those same sections listed in (A) but including the more difficult problems from the problem sets in Chapters 3 and 4.
   2. Second semester: Chapters 9 to 15, omitting 13.3 to 13.6.

II. One-Semester Course

A. Emphasis on statistics. Light coverage of Chapters 3 and 4 followed by 5 to 9 and 11. Only the rules of probability, conditional probability, independence, and expectation should be covered in Chapters 3 and 4 with the assignment of the easiest problems. Omit 7.4 to 7.7, 8.2; either 8.4 or 8.5, or both; 8.6; 9.5; and Chapter 11, except 11.1 and 11.2.

B. Emphasis on probability. A one-semester course can be given from Chapters 3 to 5 alone if one goes deeply into the problem sets and all topics.

ACKNOWLEDGMENTS

We express thanks to our editors, Leslie Hawke and Lloyd Black, for their useful ideas and substantive help. A number of colleagues have made helpful criticisms and suggestions. We name some of them below as a sign of our gratitude but emphasize that they are not responsible for any weaknesses that may remain in the text.

Mr. Edward Dowd, Mr. Pascal Diethelm, and Dr. Dev Ray of the World Health Organization have contributed significantly to our store of knowledge of survey techniques and statistical computation. Attorneys Walter Christy, Robert Glass, John Reed, and Robert Mitchell have provided us with invaluable insights into statistical applications to the Law.

Special gratitude is reserved for Professor Julia Norton of the California State University at Hayward and Professor William Notz of Purdue University who provided excellent suggestions and generous words of encouragement. Professor James Kepner of the University of Florida at Gainsville, through his trenchant reviews, greatly aided us in getting things right. We were fortunate to have other conscientious reviewers who provided incisive and constructive reviews. We take this opportunity to thank them:

William D. Bandes (San Diego Mesa College)
William E. Geeslin (University of New Hampshire at Durham)

Ken Goldstein (Miami-Dade Community College)
Michael L. Levitan (Villanova University)
Robert L. Page (University of Maine at Augusta)
Gerald E. Rubin (Marshall University)
Mark F. Schilling (University of Southern California at Los Angeles)
Donna Skane (Catonsville Community College)
Stephen A. Vardeman (Iowa State University)
Wilbur Waggoner (Central Michigan University)
I. Weiss (University of Colorado)

Our ardent thanks to Hester Paternostro, Meredith Mickel, Susan Lam, and Karen L. Bowers for their expert help joyously given in preparing the manuscript.

Contents Overview

Chapter 1	The Nature and Purposes of Statistics	1
Chapter 2	Describing Samples and Populations	11
Chapter 3	Probability Models	61
Chapter 4	Random Variables, Probability Distributions, and Probability Models	107
Chapter 5	The Normal Distribution	157
Chapter 6	Estimation; Confidence Intervals	189
Chapter 7	Hypothesis Testing About the Mean and Variance of a Population	217
Chapter 8	Comparing Two Populations	263
Chapter 9	Prediction and Association; Regression and Correlation	303
Chapter 10	Evaluation of Categorical Data; Contingency Tables	341
Chapter 11	Analysis of Variance	369
Chapter 12	Nonparametric Statistics	399
Chapter 13	Sampling Theory	437
Chapter 14	Practical Aspects of Statistics; Survey Design	469
Chapter 15	Questionnaires	495
Appendix A	Statistical Tables	527
Appendix B	Derivation of the Binomial Distribution	547
Appendix C	Confidence Belts for Proportions	549
Answers		551
Index		573

Contents

CHAPTER 1 **The Nature and Purposes of Statistics 1**

1.1 INTRODUCTION 2
1.2 WHAT STATISTICS IS 3
1.3 EXAMPLES OF DESCRIPTIVE STATISTICS 3
1.4 INFERENTIAL STATISTICS; SAMPLES, POPULATIONS 4
1.5 STEPS IN REACHING THE OBJECTIVE 6

CHAPTER 2 **Describing Samples and Populations 11**

2.1 THE VARIABLE NATURE OF DATA 12
2.2 GRAPHICAL CHARACTERIZATIONS OF DATA 13
2.3 NUMERICAL CHARACTERIZATIONS OF DATA 20
2.4 MEASURES OF CENTRAL TENDENCY 20
2.5 MEASURES OF DISPERSION OR VARIATION 33
2.6 TECHNIQUES OF DATA PRESENTATION 39
2.7 STATISTICAL COMPUTATION; CROSS-TABULATIONS 45

CHAPTER 3 **Probability Models 61**

3.1 ORIGINS OF PROBABILITY 62
3.2 INTERPRETATION OF PROBABILITY 62

3.3 MODEL OF EQUALLY LIKELY EVENTS FOR CALCULATING PROBABILITIES 64
3.4 GENERAL MODELS FOR PROBABILITIES 68
3.5 CONDITIONAL PROBABILITY AND INDEPENDENCE 72
3.6 MUTUAL INDEPENDENCE AND THE PRODUCT RULE 76
3.7 BRIDGES TO STATISTICS FROM PROBABILITY 78
3.8 SOME IMPORTANT COUNTING FORMULAS (Optional) 86

CHAPTER 4 Random Variables, Probability Distributions, and Probability Models 107

4.1 RANDOM VARIABLES 108
4.2 PROBABILITY DISTRIBUTIONS 109
4.3 PROBABILITY DISTRIBUTION AND POPULATION DISTRIBUTION 112
4.4 CONTINUOUS RANDOM VARIABLES: DENSITY FUNCTIONS AND DISTRIBUTIONS 112
4.5 MEAN AND VARIANCE OF A DISCRETE RANDOM VARIABLE 114
4.6 EXPECTATIONS 116
4.7 THE INTRODUCTION OF EXPECTED VALUE NOTATION: $E(X)$ and $V(X)$ 119
4.8 EXPECTATION OF THE SUM OF RANDOM VARIABLES 119
4.9 OTHER IMPORTANT PROPERTIES OF EXPECTATION 121
4.10 AN IMPORTANT PROPERTY OF $V(X)$ 121
4.11 DEPENDENCE AND INDEPENDENCE OF RANDOM VARIABLES: INDEPENDENT TRIALS 122
4.12 VARIANCE OF THE SUM OF MUTUALLY INDEPENDENT RANDOM VARIABLES 125

4.13 REPEATED INDEPENDENT TRIALS—THE BINOMIAL MODEL 125
4.14 OTHER MODELS (Optional) 133
4.15 THE IMPORTANCE OF THE MEAN AND VARIANCE; TCHEBYCHEV'S APPROXIMATION (Optional) 143
4.16 USING THE COMPUTER TO SIMULATE RANDOM EVENTS 144

CHAPTER 5 The Normal Distribution 157

5.1 INTRODUCTION: CALCULATION OF PROBABILITIES 158
5.2 CENTRAL LIMIT THEOREM 163
5.3 NORMAL APPROXIMATION TO THE BINOMIAL 167
5.4 HOW LARGE MUST n BE TO APPLY THE CENTRAL LIMIT THEOREM TO THE BINOMIAL? 172
5.5 SAMPLING DISTRIBUTION OF $\overline{X}$ 174
5.6 THE IMPORTANCE OF THE MEAN AND VARIANCE 178
5.7 SAMPLING EXPERIMENTS: SIMULATION OF SAMPLING FROM NORMAL POPULATIONS 178

CHAPTER 6 Estimation; Confidence Intervals 189

6.1 USING SAMPLE INFORMATION TO ESTIMATE POPULATION CHARACTERISTICS 190
6.2 UNBIASED ESTIMATES: CONFIDENCE INTERVALS ON μ WHEN σ^2 IS KNOWN 192
6.3 CONFIDENCE INTERVAL ESTIMATES FOR THE MEAN OF A NORMAL DISTRIBUTION WHEN THE POPULATION VARIANCE IS UNKNOWN 196

此为试读,需要完整PDF请访问: www.ertongbook.com

6.4 THE SAMPLING DISTRIBUTION OF S^2; CONFIDENCE INTERVALS ON THE VARIANCE OF A NORMAL DISTRIBUTION 199
6.5 THE SAMPLE PROPORTION AS AN ESTIMATE OF THE POPULATION PROPORTION 203
6.6 THE DETERMINATION OF SAMPLE SIZE 206
6.7 COMMENT ON INDEPENDENT RANDOM SAMPLING 207

CHAPTER 7 Hypothesis Testing About the Mean and Variance of a Population 217

7.1 HYPOTHESIS TESTING 218
7.2 TESTING HYPOTHESES ABOUT THE MEAN μ WHEN THE POPULATION VARIANCE σ^2 IS KNOWN 222
7.3 TESTING HYPOTHESES ABOUT THE MEAN OF A NORMAL DISTRIBUTION WHEN THE VARIANCE σ^2 IS UNKNOWN 229
7.4 TESTING HYPOTHESES ABOUT THE POPULATION VARIANCE σ^2 OF A NORMAL DISTRIBUTION 234
7.5 TEST OF HYPOTHESES CONCERNING PROPORTIONS 238
7.6 WHAT CAN GO WRONG? 244
7.7 THE POISSON DISTRIBUTION FOR RARE EVENTS. TESTS OF VERY SMALL PROBABILITIES (Optional) 246

CHAPTER 8 Comparing Two Populations 263

8.1 CONFIDENCE INTERVALS AND HYPOTHESIS TESTS COMPARING THE MEANS OF TWO POPULATIONS; VARIANCES KNOWN 264
8.2 HYPOTHESIS TESTS COMPARING THE VARIANCES OF TWO POPULATIONS; F-DISTRIBUTION 269

8.3 CONFIDENCE INTERVALS AND HYPOTHESIS TESTS ABOUT THE DIFFERENCES OF MEANS OF TWO POPULATIONS 273
8.4 CONFIDENCE INTERVALS AND TESTS OF HYPOTHESES ABOUT THE DIFFERENCE OF TWO MEANS FROM RELATED POPULATIONS 279
8.5 COMPARING TWO PROPORTIONS 283
8.6 COMPARING TWO POISSON POPULATIONS (CONTRASTING TWO RARE EVENTS) (Optional) 288
8.7 COMMENT: CONTROL AND STUDY GROUPS; THE INTERPLAY OF DATA COLLECTION AND ANALYSIS 292
8.8 COMMENTS ON COMPUTER APPLICATIONS 293

CHAPTER 9 Prediction and Association: Regression and Correlation 303

9.1 FITTING A STRAIGHT LINE 304
9.2 LINEAR REGRESSION AND THE BEST-FITTING STRAIGHT LINE 308
9.3 THE ISSUE OF ASSOCIATION: CORRELATION; THE SAMPLE CORRELATION COEFFICIENT 313
9.4 INTERPRETATION OF ρ, THE POPULATION CORRELATION COEFFICIENT 318
9.5 INFERENCE CONCERNING THE BEST-FITTING STRAIGHT LINE 323
9.6 REMARKS ON COMPUTER SOFTWARE 332

CHAPTER 10 Evaluation of Categorical Data; Contingency Tables 341

10.1 GOODNESS OF FIT 342
10.2 AN IMPORTANT EXAMPLE: TESTING FOR NORMALITY 347

10.3 TWO VARIABLES OF CLASSIFICATION—CONTINGENCY TABLES; TEST OF INDEPENDENCE 349
10.4 ASSUMPTIONS UNDERLYING THE X^2 TEST FOR CATEGORICAL DATA; A CAUTION 354
10.5 TESTS OF ASSOCIATION FOR CATEGORICAL DATA (Optional) 355
10.6 COMMENT: INTERPLAY OF DATA COLLECTION AND ANALYSIS 358

CHAPTER 11 Analysis of Variance 369

11.1 ONE-WAY ANALYSIS OF VARIANCE 370
11.2 TWO-WAY CLASSIFICATIONS WITH ONE OBSERVATION PER CELL 377
11.3 SUMMARY OF THE CHARACTERISTICS OF THE ANOVA MODEL: LINEAR MODEL 382
11.4 TWO-WAY CLASSIFICATIONS WITH MULTIPLE OBSERVATIONS PER CELL: INTERACTION 385
11.5 SENSITIVITY OF ANOVA TO THE DEPARTURES FROM THE UNDERLYING ASSUMPTIONS 389
11.6 INTERPLAY OF DATA COLLECTION AND ANALYSIS: DESIGN OF EXPERIMENTS 389

CHAPTER 12 Nonparametric Statistics 399

12.1 INTRODUCTION 400
12.2 THE SIGN TEST 401
12.3 RUNS TEST FOR RANDOMNESS 409
12.4 RANK TESTS 414
12.5 NONPARAMETRIC CONFIDENCE INTERVALS 423
12.6 COMMENTS ON NONPARAMETRIC METHODS: EFFICIENCY 425

CHAPTER 13 Sampling Theory 437

13.1 SURVEY SAMPLING 438
13.2 TERMINOLOGY 438
13.3 VARIANCE OF $\bar{X}$ FROM SIMPLE RANDOM SAMPLING 441
13.4 AN UNBIASED ESTIMATE FOR A FINITE POPULATION σ^2 442
13.5 CONFIDENCE INTERVALS 444
13.6 ESTIMATES OF PROPORTIONS 446
13.7 SAMPLE DESIGNS 446
13.8 SAMPLING AND OTHER ERRORS: METHODS FOR ACHIEVING A RANDOM SAMPLE 458
13.9 REMARKS 461

CHAPTER 14 Practical Aspects of Statistics; Survey Design 469

14.1 INTRODUCTION 470
14.2 CENSUS OR SAMPLE? 470
14.3 HISTORICAL SUMMARY 471
14.4 STAGES IN A SAMPLE SURVEY 472
14.5 TYPES OF SURVEYS 480
14.6 CHECKING INTERVIEWERS: WHO'S LYING, WHO'S UNPOPULAR 482
14.7 APPLICATION: THE BREAST-FEEDING CONTROVERSY (Optional) 483

CHAPTER 15 Questionnaires 495

15.1 INTRODUCTION 496
15.2 THE STRUCTURE OF A QUESTIONNAIRE 496
15.3 NONSAMPLING ERRORS 498
15.4 RELIABILITY AND VALIDITY 501
15.5 GENERAL PURPOSES OF QUESTIONNAIRES; ANALYSIS OF CATEGORICAL QUESTIONNAIRES 510

15.6 INTERNAL CONSISTENCY—EVALUATING EACH QUESTION FOR WORTH IN CATEGORICAL QUESTIONNAIRES 514
15.7 DISCRIMINATION CATEGORICAL QUESTIONNAIRES (Optional) 517
15.8 A FINAL WORD 520

APPENDIX A Statistical Tables 527

A.1 NORMAL CURVE AREAS AND COMMON UPPER CENTILES 528
A.2 COMBINATION VALUES 530
A.3 TABLE OF RANDOM NUMBERS 531
A.4 TABULATED VALUES OF e^{-x} 533
A.5 t-DISTRIBUTION UPPER CENTILES, t_α 534
A.6 UPPER CENTILE VALUES X^2_α 536
A.7 UPPER CENTILES F_α, (df(N), df(D)) 538
A.8 BINOMIAL PROBABILITIES 540
A.9 VALUES OF $F(v) = \frac{1}{2}ln(1+v)/(1-v)$ 541
A.10 DISTRIBUTION OF TOTAL NUMBER OF RUNS u IN SAMPLES OF SIZE (n_1, n_2) 542
A.11 DISTRIBUTION OF THE SIGNED-RANK STATISTIC W 544
A.12 PERCENTILE OF r WHEN $\rho = 0$ 545
A.13 PERCENTILES OF THE MAXIMUM DEVIATION STATISTICS 545
A.14 CONFIDENCE INTERVALS FOR THE MEDIAN 546

APPENDIX B Derivation of the Binomial Distribution 547

APPENDIX C Confidence Belts for Proportions 549

Answers 551

Index 573

CHAPTER • 1

The Nature and Purposes of Statistics

"Let us sit on this log at the roadside, says I, and forget the inhumanity and the ribaldry of the poets. It is in the glorious columns of ascertained facts and legalized measures that beauty is to be found. In this very log that we sit upon, Mrs. Sampson, says I, is statistics more wonderful than any poem."

O. Henry
The Handbook of Hymen

Statistical information plays an important part in a wide range of fields, many of which intimately affect our lives. In this chapter we study two main areas in statistics: descriptive and inferential statistics. We also explore two concepts basic to all our subsequent discussions: the "population" and the "sample."

此为试读,需要完整PDF请访问: www.ertongbook.com