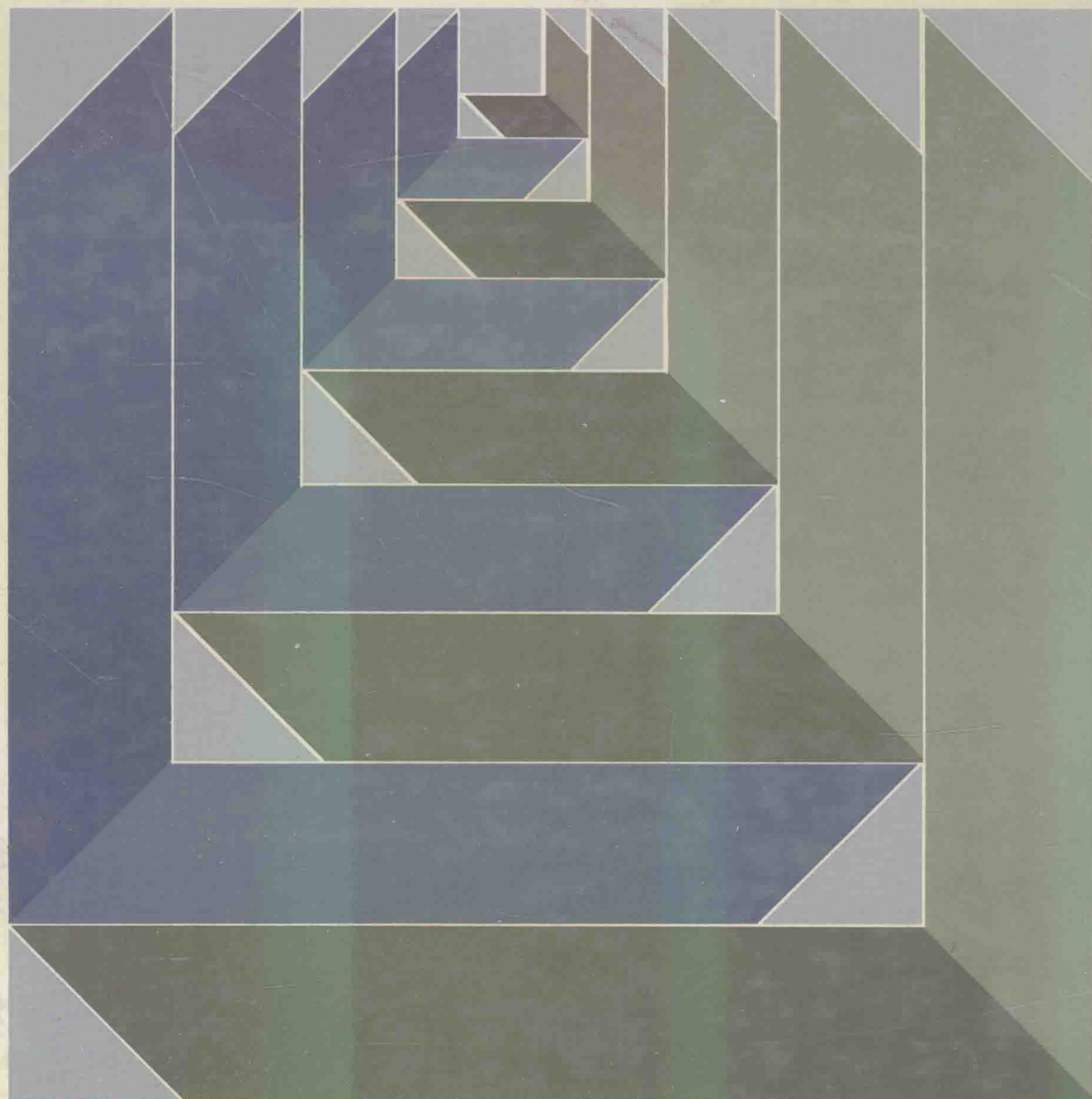


GENE V GLASS / KENNETH D. HOPKINS

STATISTICAL METHODS IN EDUCATION AND PSYCHOLOGY



SECOND EDITION

STATISTICAL METHODS IN EDUCATION AND PSYCHOLOGY

second edition

GENE V GLASS

*College of Education
Arizona State University, Tempe*

KENNETH D. HOPKINS

*Laboratory of Educational Research
University of Colorado*

PRENTICE-HALL, INC., *Englewood Cliffs, New Jersey 07632*

Library of Congress Cataloging in Publication Data

Glass, Gene V

Statistical methods in education and psychology.

Bibliography: p.

Includes index.

1. Education statistics. 2. Psychometrics.

I. Hopkins, Kenneth D. II. Title.

LB2846. G55 1984 519 84-1954

ISBN 0-13-844944-9

Editorial/production supervision and

interior design: Marion Osterberg and Fred Bernardi

Cover design: Wanda Lubelska Design

Manufacturing buyer: Ron Chapman

*To the graduates of the Laboratory of Educational Research,
University of Colorado-Boulder, who teach and apply statistics
all over the world, for their contributions to our thinking about statistical methods.*

© 1984, 1970 by Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632

*All rights reserved. No part of this book may be
reproduced, in any form or by any means,
without permission in writing from the publisher.*

Printed in the United States of America

10 9 8 7 6

ISBN 0-13-844944-9 01

Prentice-Hall International, Inc., London

Prentice-Hall of Australia Pty. Limited, Sydney

Editora Prentice-Hall do Brasil, Ltda., Rio de Janeiro

Prentice-Hall Canada Inc., Toronto

Prentice-Hall of India Private Limited, New Delhi

Prentice-Hall of Japan, Inc., Tokyo

Prentice-Hall of Southeast Asia Pte. Ltd., Singapore

Whitehall Books Limited, Wellington, New Zealand

PREFACE

This textbook is designed for a one- or two-semester course in applied statistics. The methods are applicable to empirical research in many disciplines. We have drawn applications from several disciplines, although most come from education and psychology. In most instances, the data are not hypothetical but are from actual studies.

Our selection of topics has been guided by three considerations: (1) What are the most useful statistical methods? (2) Which statistical techniques are the most widely used in scholarly journals in the behavioral and social sciences? (3) Which statistical concepts and methods are fundamental to further study?

This text attempts to produce a more thorough coverage of analysis of variance techniques than usually appears in basic statistics texts for social scientists. Edgington (1974) found that more than 70 percent of all articles published in journals of the American Psychological Association employed this extremely versatile and powerful technique. According to the survey of Wick and Dirkes (1973), more than 95 percent of the doctoral dissertations in education use the statistical techniques that are treated in this text.

This edition of *Statistical Methods in Education and Psychology (SMEP)* differs from the first edition in both substance and style. The substantive changes include (1) an expanded and integrated treatment of multiple comparisons, including trend analysis, (2) a new chapter on multiple regression, (3) a new chapter on inferential methods with proportions, and (4) expanded treatment of more complex applications of the analysis of variance and covariance.

The pedagogical improvements include (1) diagnostic mastery tests following each chapter, (2) reorganization of content involving separate chapters on inferences regarding means, variances, proportions, and correlation coefficients, and (3) an attempt to provide more options in course content and emphasis. *SMEP* has endeavored to continue to be “extraordinary in its collection of examples, exercises, and statistical tables” (Schmidt, 1972, p. 169).

The approach of this text is conceptual, not mathematical or “cookbookish.” Deriving a formula is no proof of real understanding, nor is the ability to plug numbers into formulas and “turn the crank.” Indeed, the number of formulas used is kept small; the verbal-to-mathematical ratio of text material would rank high among statistics texts. We have stressed concepts rather than derivation and proof.

We have pruned away much deadwood present in some other statistics texts. Although every text claims to reflect the latest influences of electronic computers and hand calculators, obsolete techniques of calculating the mean, variance, and correlation coefficient from “grouped” data continue to be taught. These “shortcut” methods lead to roundabout formulas that, in addition to being less accurate, impede conceptualizing the meaning of the statistic being calculated.

We have tried to be sensitive to changes in statistical pedagogy occasioned by the rapid spread of hand calculators. We strongly advise students to purchase a hand calculator having at least one memory and the square root (\sqrt{x}) and reciprocal ($1/X$) features; better yet would be one with preprogrammed functions for mean, standard deviation, and correlation.

In this edition, we have tried to maintain the high professional standards of the first edition of *SMEP*. Michael (1970, pp. 1015 and 1018) found that the first edition of *SMEP* “probably affords the most nearly current and best balanced treatment of statistical methodology to be found in any well-known introductory text. . . . This volume should set the standard for many years to come.”

In *Twenty-Five Years of Recommended Readings in Psychology*, Solso (1979) found that *SMEP* was among the 25 most recommended books by graduate departments of psychology, a distinction recorded by only one other statistics book (and it was less popular than *SMEP*). In their survey of statistics professors in graduate departments of education, Gay, Campbell, and Gallagher (1978) found that *SMEP*, in comparison to the other common statistics texts studied, had the highest ratings of (1) technical accuracy, (2) breadth of coverage, and (3) depth of topic coverage. *SMEP* was judged (1) to have no statistics prerequisites for comprehension, (2) to have a good sequence of topics, and (3) to be appropriate for undergraduate and graduate work, and (4) to have an orientation that is appropriate for both research consumers and producers. Stein and Kuenne (1979) found *SMEP* to have the lowest (easiest) readability of the seven statistics and research textbooks they studied. In a recent survey of the 100 top research-producing universities in education, *SMEP* was the most-used text (Brinzer & Sinatra, 1982). During the two years this text was being revised, our colleagues and students contributed in innumerable ways to our efforts. We cannot name them all here, but the following colleagues deserve special thanks for contributing recently to our education in statistical methods: Julian Stanley deserves credit for his influence as mentor and his contributions to the first edition. His modesty and heavy research commitments led to his decision not to remain a co-author for the second edition. Roberta Flexer, Stuart Kahl, Jason Millman, and Lorrie Shepard made several suggestions based on extensive classroom experience with *SMEP* that improved it as a text. Others made technical or pedagogical contributions: Maurice Tatsuoka, Edward Cureton, William Michael, Lorrie Shepard, James Collins,

Roberta Flexer, Stuart Kahl, Victor Willson, Todd Rogers, Arlen Gullickson, Percy Peckham, Bob Hopkins, Mary Lee Smith, James Sanders, Stephen Jurs, Gregory Camilli, George Kretke, James Morrow, Frank B. Baker, and Carol Vojir.

We are indebted to Bob Hopkins, Mark Chisholm, James Rosenthal, Rick Kroc, Thomas Hughes, Lynn Sherretz, Alan Davis, Paul Kelter, and Carol Vojir for critically reading various chapters and providing answers to problems and exercises.

The tedious tasks of deciphering, typing, and correcting various drafts of the manuscript were shared by Sue Reissig, Karen O'Grady, Mary Ann Gardner, and Sheryl Vriesman.

Gene V Glass
Kenneth D. Hopkins

CONTENTS

PREFACE xi

1 INTRODUCTION 1

1.1 The “Image” of Statistics, 1 1.2 Descriptive Statistics, 2 1.3 Inferential Statistics, 2 1.4 Statistics and Mathematics, 3

2 VARIABLES, MEASUREMENT, AND SCALES 5

2.1 Variables and Their Measurement, 5 2.2 Measurement: The Observation of Variables, 6 2.3 Measurement Scales: Nominal Measurement, 6 2.4 Ordinal Measurement, 7 2.5 Interval Measurement, 7 2.6 Ratio Measurement, 8 2.7 Interrelationships Among Measurement Scales, 9 2.8 Continuous and Discrete Variables, 9 Chapter Summary, 10 Mastery Test, 10 Mastery Test Answers, 11

3 FREQUENCY DISTRIBUTIONS 12

3.1 Tabulation of Data, 12 3.2 Rank-Order Distributions, 12 3.3 The Ungrouped Frequency Distribution, 13 3.4 Grouped Frequency Distributions, 15 3.5 Grouping

and Loss of Information, 17 3.6 Graphing a Frequency Distribution: The Histogram, 17 3.7 Frequency and Percentage Polygons, 18 3.8 Describing Distributions, 20 3.9 Cumulative Distributions, 21 3.10 Percentiles, 23 3.11 Box-and-Whisker Plots, 23 3.12 Stem-and-Leaf Displays, 24 3.13 Time-Series Graphs, 26 Chapter Summary, 26 Mastery Test, 27 Problems and Exercises, 28 Answers to Mastery Test, 30 Answers to Problems and Exercises, 30

4 MEASURES OF CENTRAL TENDENCY 31

4.1 Introduction, 31 4.2 The Mode, 31 4.3 The Median, 32 4.4 Summation Notation, Σ , 33 4.5 The Mean, 34 4.6 More Summation Notation, 35 4.7 Adding or Subtracting a Constant, 36 4.8 Multiplying or Dividing by a Constant, 36 4.9 Sum of Deviations, 36 4.10 Sum of Squared Deviations, 37 4.11 The Mean of the Sum of Two or More Scores, 37 4.12 The Mean of a Difference, 38 4.13 Mean, Median, and Mode of Two or More Groups, 38 4.14 Interpretation of Mode, Median, and Mean, 40 4.15 Other Measures of Central Tendency, 40 4.16 Central Tendency and Skewness, 41 4.17 Measures of Central Tendency as Inferential Statistics, 42 4.18 Which Measure is Best?, 42 Chapter Summary, 43 Mastery Test, 43 Problems and Exercises, 44 Answers to Mastery Test, 45 Answers to Problems and Exercises, 45

5 MEASURES OF VARIABILITY 46

5.1 Introduction, 46 5.2 The Range, 46 5.3 The H-Spread and the Interquartile Range, 47 5.4 Deviation Scores, 47 5.5 Sum of Squares, 48 5.6 More About the Summation Operator, Σ , 48 5.7 The Variance of a Population, 49 5.8 The Variance Estimated from a Sample, 49 5.9 The Standard Deviation, 51 5.10 Adding or Subtracting a Constant, 52 5.11 Multiplying or Dividing by a Constant, 52 5.12 Variance of a Composite Distribution, 53 5.13 Inferential Properties of the Range, s^2 , and s , 54 Chapter Summary, 55 Mastery Test, 56 Problems and Exercises, 57 Answers to Mastery Test, 57 Answers to Problems and Exercises, 58

6 THE NORMAL DISTRIBUTION AND STANDARD SCORES 59

6.1 The Importance of the Normal Distribution, 59 6.2 The Normal Curve, 62 6.3 The Unit-Normal Distribution as a Standard: z-Scores, 62 6.4 Ordinates of the Unit-Normal Distribution, 63 6.5 Areas Under the Normal Curve, 64 6.6 Other Standard Scores, 64 6.7 Areas Under Normal Curve in Samples, 67 6.8 Skewness, 68 6.9 Kurtosis, 70 6.10 Transformations, 70 6.11 Normalized Scores, 72 Chapter Summary, 73 Mastery Test, 73 Problems and Exercises, 74 Answers to Mastery Test, 77 Answers to Problems and Exercises, 77

7 CORRELATION: THE MEASUREMENT OF RELATIONSHIP 79

7.1 Introduction, 79 7.2 The Concept of Correlation, 80 7.3 The Measurement of Correlation, 80 7.4 Scatterplots, 81 7.5 Linear and Curvilinear Relationships, 83

7.6 Calculating the Pearson Product-Moment Correlation Coefficient, r , 84 7.7 A Computational Illustration of r , 85 7.8 Alternative Computational Formulas for r , 88 7.9 Linear Transformations and Correlation, 89 7.10 The Bivariate Normal Distribution, 90 7.11 The Pearson r and Marginal Distributions, 91 7.12 Effects of Variability on Correlation, 92 7.13 The Variance of a Sum, 93 7.14 The Variance of Difference Scores, 94 7.15 Additional Measures of Relationship: The Spearman Rank Correlation, 95 7.16 The Phi Coefficient: Both X and Y are Dichotomies, 97 7.17 The Point-Biserial Coefficient, 100 7.18 The Biserial Correlation, 100 7.19 Biserial Versus Point-Biserial Correlation Coefficients, 102 7.20 The Tetrachoric Coefficient, 102 7.21 The Rank-Biserial Correlation Coefficient, 103 7.22 Causation and Correlation, 104 Chapter Summary, 107 Mastery Test, 107 Problems and Exercises, 110 Answers to Mastery Test, 112 Answers to Problems and Exercises, 113

8 LINEAR AND MULTIPLE REGRESSION 114

8.1 Purposes of Regression Analysis, 114 8.2 The Regression Effect, 115 8.3 The Regression Equation Expressed in Standard z -Scores, 116 8.4 Use of Regression Equations, 117 8.5 Cartesian Coordinates, 117 8.6 Estimating Y from X : The Raw-Score Regression Equation, 119 8.7 Error of Estimate, 121 8.8 Proportion of Predictable Variance, r^2 , 121 8.9 Least-Squares Criterion, 122 8.10 Homoscedasticity and the Standard Error of Estimate, 123 8.11 Regression and Pretest-Posttest Gains, 127 8.12 Part Correlation and Partial Correlation, 128 8.13 Partial Correlation, 130 8.14 Second-Order Partial Correlations, 131 8.15 Multiple Regression and Multiple Correlation, 131 8.16 The Standardized Regression Equation, 132 8.17 The Raw-Score Regression Equation, 134 8.18 Multiple Correlation, 135 8.19 Multiple Regression Equations with Three or More Independent Variables, 136 8.20 Stepwise Multiple Regression, 136 8.21 An Illustration of Stepwise Multiple Regression, 137 8.22 Dichotomous and Categorical Variables as Predictors, 138 8.23 The Standard Error of Estimate in Multiple Regression, 139 8.24 The Multiple Correlation as an Inferential Statistic: Correction for Bias, 139 8.25 Assumptions, 141 8.26 Curvilinear Regression and Correlation, 141 8.27 Measuring Nonlinear Relationships Between Two Variables: The Correlation Ratio η^2 , 141 8.28 Transforming Nonlinear Relationships into Linear Relationships, 144 Chapter Summary, 144 Mastery Test, 144 Problems and Exercises, 148 Answers to Mastery Test, 151 Answers to Problems and Exercises, 152

9 PROBABILITY 154

9.1 Introduction, 154 9.2 Probability as a Mathematical System, 155 9.3 First Addition Rule of Probabilities, 156 9.4 Second Addition Rule of Probabilities, 158 9.5 Multiplicative Rule of Probabilities, 160 9.6 Conditional Probability, 161 9.7 Bayes' Theorem, 161 9.8 Permutations, 162 9.9 Combinations, 163 9.10 Intuition and Probability, 164 9.11 Expectations and Moments, 167 Chapter Summary, 168 Mastery Test, 169 Problems and Exercises, 170 Answers to Mastery Test, 171 Answers to Problems and Exercises, 171

10 STATISTICAL INFERENCE: SAMPLING AND INTERVAL ESTIMATION 173

10.1 The Purpose of Statistical Inference, 173 10.2 Populations and Samples: Parameters and Statistics, 174 10.3 Infinite Versus Finite Populations, 174 10.4 Randomness and Random Sampling, 175 10.5 Accidental or Convenience Samples, 176 10.6 Random Samples, 176 10.7 Independence, 178 10.8 Systematic Sampling, 179 10.9 Point and Interval Estimates, 180 10.10 Sampling Distributions, 180 10.11 The Standard Error of the Mean When X Is Normally Distributed, 180 10.12 Relationship of $\sigma_{\bar{x}}$ to n , 181 10.13 Confidence Intervals, 182 10.14 Confidence Intervals When σ Is Known: An Example, 183 10.15 Central Limit Theorem, 184 10.16 The Use of Sampling Distributions, 188 10.17 Proof that $\sigma_{\bar{x}}^2 = \sigma^2/n$, 188 10.18 Properties of Estimators, 191 10.19 Unbiasedness, 191 10.20 Consistency, 194 10.21 Relative Efficiency, 194 Chapter Summary, 196 Mastery Test, 197 Problems and Exercises, 199 Answers to Mastery Test, 200 Answers to Problems and Exercises, 200

11 INTRODUCTION TO HYPOTHESIS TESTING 202

11.1 Introduction, 202 11.2 Statistical Hypotheses and Explanations, 203 11.3 Statistical Versus Scientific Hypotheses, 203 11.4 Testing Statistical Hypotheses About μ , 204 11.5 Testing $H_0: \mu = K$, a One-Sample z -Test, 205 11.6 Types of Errors in Hypothesis Testing, 206 11.7 Hypothesis Testing and Confidence Intervals, 208 11.8 Type-II Error, β , and Power, 209 11.9 Power, 210 11.10 Effect of α on Power, 211 11.11 Power and the Value Hypothesized in the Alternative Hypothesis, 212 11.12 Methods of Increasing Power, 213 11.13 Nondirectional and Directional Alternatives: Two-Tailed Versus One-Tailed Tests, 214 11.14 Statistical Significance Versus Practical Significance, 215 11.15 Confidence Limits for the Population Median, 216 11.16 Inferences Regarding μ When σ Is Not Known: t Versus z , 217 11.17 The t -Distribution, 218 11.18 Confidence Intervals Using the t -Distribution, 220 11.19 Accuracy of Confidence Intervals When Sampling Nonnormal Distributions, 221 Chapter Summary, 222 Mastery Test, 223 Problems and Exercises, 224 Answers to Mastery Test, 226 Answers to Problems and Exercises, 227

12 INFERENCES ABOUT THE DIFFERENCE BETWEEN MEANS 229

12.1 Introduction, 229 12.2 Testing Statistical Hypotheses Involving Two Means, 230 12.3 The Null Hypothesis, $H_0: \mu_1 - \mu_2 = 0$, 230 12.4 The t -Test Comparing Two Independent Means, 230 12.5 Computing $s_{\bar{x}_1 - \bar{x}_2}$, 232 12.6 An Illustration, 233 12.7 Confidence Intervals About Mean Differences, 235 12.8 Effect Size, 236 12.9 Normality, 237 12.10 Homogeneity of Variance, 238 12.11 Independence of Observations, 240 12.12 Testing $H_0: \mu_1 = \mu_2$ With Paired Observations, 240 12.13 Direct-Difference Method for the t -Test With Dependent Observations, 243 12.14 Cautions Regarding the Matched-Pair Designs in Research, 245 Chapter Summary, 246 Mastery Test, 246 Problems and Exercises, 249 Answers to Mastery Test, 253 Answers to Problems and Exercises, 253

13 INFERENCE REGARDING VARIANCES 255

13.1 Introduction, 255 13.2 Chi-Square Distributions, 256 13.3 The Chi-Square Distribution with ν Degrees of Freedom, χ^2_ν , 258 13.4 Inferences About the Population Variance: $H_0: \sigma^2 = K$, 259 13.5 F -Distributions, 261 13.6 Inferences About Two Independent Variances: $H_0: \sigma_1^2 = \sigma_2^2$, 262 13.7 Testing Homogeneity of Variance: Hartley's F_{max} Test, 265 13.8 Testing Homogeneity of Variance from J Independent Samples: The Bartlett Test, 266 13.9 Inferences About σ_1^2/σ_2^2 with Paired Observations, 268 13.10 Relationships Among the Normal, t -, Chi-Square, and F -Distributions, 269 Chapter Summary, 270 Mastery Test, 271 Problems and Exercises, 272 Answers to Mastery Test, 273 Answers to Problems and Exercises, 274

14 INFERENCES REGARDING PROPORTIONS 275

14.1 Statistics for Categorical Dependent Variables, 275 14.2 The Proportion as a Mean, 276 14.3 The Variance of a Proportion, 276 14.4 The Standard Error and Sampling Distribution of a Proportion, 277 14.5 Testing $H_0: \pi = K$, 279 14.6 Testing Empirical Versus Theoretical Distributions: The Chi-Square Goodness-of-Fit Test, 282 14.7 Test of Normality, 284 14.8 Testing Differences in Proportions: The Chi-Square Test of Association, 285 14.9 The χ^2 Median Test, 289 14.10 Chi-Square and the Phi Coefficient, 290 14.11 Independence of Observations, 291 14.12 Inferences About $\pi_1 = \pi_2$ When Observations Are Paired, 291 Chapter Summary, 293 Mastery Test, 293 Problems and Exercises, 295 Answers to Mastery Test, 298 Answers to Problems and Exercises, 298

15 INFERENCES AMONG CORRELATION COEFFICIENTS 300

15.1 Testing a Statistical Hypothesis Regarding ρ , 300 15.2 Testing $H_0: \rho = 0$ Using the t -Test, 301 15.3 Directional Alternatives: "Two-Tailed" Versus "One-Tailed" Tests, 303 15.4 Setting Confidence Intervals About r : The Fisher Z -Transformation, 304 15.5 Determining Confidence Intervals Graphically, 307 15.6 Testing the Difference Between Correlation Coefficients: $H_0: \rho_1 = \rho_2$, 307 15.7 Testing Differences Among Several Independent Correlation Coefficients: $H_0: \rho_1 = \rho_2 = \dots = \rho_J$, 309 15.8 Testing Differences Between Two Dependent Correlation Coefficients: $H_0: \rho_{31} = \rho_{32}$, 310 15.9 Inferences About Other Correlation Coefficients, 311 15.10 The Point-Biserial Correlation Coefficient r_{pb} , 312 15.11 Spearman's Rank Correlation: $H_0: \rho_{\text{ranks}} = 0$, 313 15.12 Partial Correlation, $H_0: \rho_{12.3} = 0$, 313 15.13 Significance of a Multiple Correlation Coefficient, 314 15.14 Statistical Significance in Stepwise Multiple Regression, 314 15.15 Significance of the Biserial Correlation Coefficient r_{bis} , 315 15.16 Significance of the Tetrachoric Correlation Coefficient r_{ter} , 316 15.17 Significance of the Correlation Ratio, 317 15.18 Testing for Nonlinearity of Regression, 318 Chapter Summary, 318 Mastery Test, 319 Problems and Exercises, 320 Answers to Mastery Test, 322 Answers to Problems and Exercises, 322

16 THE ONE-FACTOR ANALYSIS OF VARIANCE 324

16.1 Introduction, 324 16.2 Why Not Several t -Tests?, 324 16.3 ANOVA Nomenclature, 326 16.4 ANOVA Computation, 326 16.5 ANOVA Computational Illustration, 329 16.6 ANOVA Theory, 330 16.7 The F -Test, 333 16.8 A Model for the Data, 334 16.9 Estimates of the Terms in the Model, 335 16.10 Sums of Squares, 336 16.11 Restatement of the Null Hypothesis in Terms of Population Means, 337 16.12 Degrees of Freedom, 337 16.13 Mean Squares: Expectation of MS_b and MS_w , 339 16.14 Some Distribution Theory, 341 16.15 The F -Test of the Null Hypothesis: Rationale and Procedure, 344 16.16 Type-I Versus Type-II Errors: α and β , 346 16.17 Computational Procedures for the One-Factor ANOVA, 348 16.18 Consequences of Failure to Meet the ANOVA Assumptions: The "Robustness" of ANOVA, 350 16.19 The Power of the F -Test, 353 16.20 An Illustration, 354 16.21 Power When σ Is Unknown, 355 16.22 n Required for Specified Power, 355 16.23 The Scheffé Test for Homogeneity of Variance, 357 Chapter Summary, 359 Mastery Test, 359 Problems and Exercises, 362 Answers to Mastery Test, 365 Answers to Problems and Exercises, 366

17 MULTIPLE COMPARISONS 368

17.1 Introduction, 368 17.2 Testing All Pairs of Means: The Studentized Range Statistic, q , 370 17.3 The Tukey Method of Multiple Comparisons, 371 17.4 The Effect Size of Mean Differences, 373 17.5 The Basis for Type-I Error Rate: Contrast Versus Family, 373 17.6 The Newman-Keuls Method, 374 17.7 The Tukey and Newman-Keuls Methods Compared, 376 17.8 The Definition of a Contrast, 376 17.9 Simple Versus Complex Contrasts, 377 17.10 The Standard Error of a Contrast, 379 17.11 The t -Ratio for a Contrast, 380 17.12 Planned Versus Post Hoc Comparisons, 380 17.13 Dunn Method of Multiple Comparisons, 381 17.14 Dunnett Method of Multiple Comparisons, 381 17.15 Scheffé Method of Multiple Comparisons, 382 17.16 Planned Orthogonal Contrasts, 383 17.17 Confidence Intervals for Contrasts, 384 17.18 Relative Power of Multiple Comparison Techniques, 386 17.19 Trend Analysis, 386 17.20 Significance of Trend Components, 389 17.21 Relations of Trends to Correlation Coefficients, 390 17.22 Assumptions of MC Methods and Trend Analysis, 391 17.23 Multiple Comparisons Among Other Statistics, 391 Chapter Summary and Criteria for Selecting a Multiple Comparison Method, 392 Mastery Test, 394 Problems, 397 Exercises, 398 Answers to Mastery Test, 400 Answers to Problems, 400 Answers to Exercises, 401

18 TWO- AND THREE-FACTOR ANALYSIS OF VARIANCE: An Introduction to Factorial Designs 402

18.1 Introduction, 402 18.2 The Meaning of Interaction, 403 18.3 Interaction and Generalization, 404 18.4 Interpreting Main Effects When Interaction Is Present, 406 18.5 Two Types of Interaction, 408 18.6 Data Layout and Notation, 409 18.7 A Model for the Data, 411 18.8 Least-Squares Estimation of the Model, 412 18.9 State-

ment of Null Hypotheses, 414 18.10 Sums of Squares in the Two-Factor ANOVA, 417
 18.11 Degrees of Freedom, 419 18.12 Mean Squares, 421 18.13 Illustration of
 Computation for the Two-Factor ANOVA, 421 18.14 Expected Values of Mean
 Squares, 424 18.15 The Distributions of the Mean Squares, 428 18.16 Hypothesis
 Tests of the Null Hypotheses, 431 18.17 Determining Power in Factorial Designs, 433
 18.18 Multiple Comparisons in the Two-Factor ANOVA, 434 18.19 Confidence
 Intervals for Means in Two-Factor ANOVA, 436 18.20 Three-Factor ANOVA, 437
 18.21 Three-Factor ANOVA: An Illustration, 437 18.22 Three-Factor ANOVA Com-
 putation, 439 18.23 The Interpretation of Three-Factor Interactions, 441 18.24 Con-
 fidence Intervals in Three-Factor ANOVA, 441 18.25 Multiple Comparisons with
 Three-Factor ANOVA, 442 18.26 Factorial Designs and Power, 443 18.27 Factorial
 ANOVA with Unequal n 's, 444 Chapter Summary, 444 Mastery Test, 445 Problems
 and Exercises, 447 Answers to Mastery Test, 451 Answers to Problems, 452

19 MULTI-FACTOR ANALYSIS OF VARIANCE: Random, Mixed, and Fixed Effects 456

19.1 Introduction, 456 19.2 The Random-Effects ANOVA Model, 456 19.3 As-
 sumptions of the Random ANOVA Model, 458 19.4 An Example, 459 19.5 Mean
 Square Within, 460 19.6 Mean Square Between, 461 19.7 The Variance Component
 σ_a^2 , 462 19.8 Confidence Interval for σ_a^2/σ_e^2 , 463 19.9 Summary of Random ANOVA
 Model, 465 19.10 The Mixed-Effects ANOVA Model, 465 19.11 Mixed-Model
 ANOVA Assumptions, 468 19.12 Mixed-Model ANOVA Computation, 469 19.13 Mul-
 tiple Comparisons in the Two-Factor Mixed Model, 471 19.14 Repeated-Measures De-
 signs, 472 19.15 Crossed and Nested Factors, 476 19.16 Computation of Sums of
 Squares for Nested Factors, 477 19.17 Determining the Sources of Variation: The
 ANOVA Table, 478 19.18 Degrees of Freedom for Nested Factors, 478 19.19 Deter-
 mining Expected Mean Squares, 479 19.20 Error Mean Square in Complex ANOVA
 Designs, 480 19.21 Model Simplification and Pooling, 481 19.22 The Experimental
 Unit and the Observational Unit, 481 Chapter Summary, 484 Mastery Test, 484
 Problems and Exercises, 487 Answers to Mastery Test, 490 Answers to Problems
 and Exercises, 490

20 AN INTRODUCTION TO THE ANALYSIS OF COVARIANCE 492

20.1 The Functions of ANCOVA, 492 20.2 ANOVA Results, 493 20.3 ANCOVA
 Model, 494 20.4 ANCOVA Computations: SS_{total} , 496 20.5 The Adjusted Within
 Sum of Squares SS'_w , 497 20.6 The Adjusted Sum of Squares Between Groups SS'_b , 498
 20.7 Degrees of Freedom in ANCOVA and the ANCOVA Table, 498 20.8 Adjusted
 Means, Y'_j , 499 20.9 Confidence Intervals and Multiple Comparisons for Adjusted
 Means, 501 20.10 ANCOVA Illustrated Graphically, 501 20.11 ANCOVA Assump-
 tions, 503 20.12 ANCOVA Precautions, 505 20.13 Covarying Versus Stratifying, 506
 Chapter Summary, 507 Mastery Test, 507 Problems and Exercises, 508 Answers to
 Mastery Test, 510 Answers to Problems and Exercises, 510

APPENDIX 512

Rules of Thumb for Writing the ANOVA Table, 512 I. Definitions of Terms, 513
II. Determining the Possible Lines (Sources of Variation) of the ANOVA Table, 514
III. Determining the Degrees of Freedom for Sources of Variation, 515 IV. Computing
Sums of Squares, 516 V. Determining the Expectations of Mean Squares, 518

TABLES 521

Table A: Unit-Normal (z -) Distribution, 522 Table B: Random Digits, 528 Table C:
 t -Distribution, 529 Table D: χ^2 -Distribution, 531 Table E: Fisher's Z -Transformation, 533
Table F: Critical Values of F , 534 Table G: Power Curves for the F -Test, 542
Table H: Hartley's F_{\max} -Distribution, 545 Table I: Studentized Range Statistic: q -Distribution, 546
Table J: Critical Value of r , 549 Table K: Critical Values of r_{ranks} ,
Spearman's Rank Correlation, 550 Table L: Critical Values for the Dunn (Bonferroni)
 t -Statistic, 551 Table M: Critical t -Values for the Dunnett t -Statistic, 554 Table N:
Coefficients (Orthogonal Polynomials) for Trend Analysis, 556

BIBLIOGRAPHY 557**AUTHOR INDEX 567****SUBJECT INDEX 570**

I

INTRODUCTION

1.1 THE “IMAGE” OF STATISTICS

Popular attitudes toward statistics contain a mixture of awe, cynicism, suspicion, and contempt. Freudian slips have transformed statistics into “sadistics,” and “Don’t become a statistic” is taken to mean “Don’t let something evil befall you.”

Statisticians have been scornfully placed in the company of liars and accused of “statisticulation”—the art of lying with statistics while maintaining an appearance of objectivity and rationality. Someone once remarked: “If all the statisticians in the world were laid end to end—it would be a good thing.” A statistician has been depicted as a person who drowns while wading in a river having an average depth of 3 ft., or who sits with his head in a refrigerator and his feet in an oven and reports, “On the average, I feel fine.”

W. H. Auden wrote, “Thou shalt not sit among statisticians, nor commit a social science.” But nonsense can be expressed as readily verbally as it can be quantitatively. Knowledge of logic is a good safeguard against uncritical acceptance of verbal nonsense, and knowledge of statistics is the best defense against quantitative nonsense. The study of statistical concepts and methods will certainly reduce numerical credulity and help one be a wise consumer of quantitative information. The first step toward replacing popular

images of statistics with more realistic ones is to study the structure of the discipline of “statistical methods” and its historical antecedents.

Some persons avoid statistics because of philosophical bias, apprehension about its rigors, or misconceptions about the discipline. Some prefer to operate on the basis of tradition, intuition, authoritative judgment, or “common sense.” But it is increasingly recognized that there is a place for systematic, objective, and empirical research for which statistics is a tool.

Knowledge of statistical methods is becoming necessary for scholarship in most empirical disciplines. In the past twenty years, most graduate schools have acknowledged its importance as a research tool by accepting course work in statistics as a substitute for one of the two foreign languages traditionally required for a Ph.D. degree. The substitution is apt: statistics is an increasingly important means of communicating knowledge.

There were two widely divergent influences on the early development of statistical methods. Statistics had a mother who was dedicated to keeping orderly records of governmental units (*state* and *statistics* come from the same Latin root, *status*) and a gambling father who relied on mathematics to increase his skill at playing the odds in games of chance. From the mother sprang counting, measuring, describing, tabulating, ordering, and the taking of censuses—all of which led to modern *descriptive statistics*. From the father eventually came modern *inferential statistics*, which is based squarely on theories of probability. This text offers an introduction to the descriptive and inferential statistics that are most widely used in educational and behavioral research (Edgington, 1974; Wick and Dirkes, 1973; Willson, 1980). Descriptive statistics are emphasized in Chapters 2 through 8. Beginning with “Probability” in Chapter 9 and extending through Chapter 15, topics from inferential statistics are covered. Chapters 16 through 20 present the considerations and inferential techniques especially important in the design and analysis of experiments.

1.2

DESCRIPTIVE STATISTICS

Descriptive statistics involves tabulating, depicting, and describing collections of data. These data may be either *quantitative*, such as measures of height or intelligence—variables that are characterized by an underlying continuum—or the data may represent *qualitative* variables, such as sex, college major, or personality type. Large masses of data generally must undergo a process of summarization or reduction before they are comprehensible.

The human mind cannot extract the full import of a mass of data (How do they vary? About how large are they? Is one set useless in reducing uncertainty about the other?) without the aid of special techniques. Thus descriptive statistics serves as a tool to describe or summarize or reduce to manageable form the properties of an otherwise unwieldy mass of data.

1.3

INFERENTIAL STATISTICS

Inferential statistics is a formalized body of methods for solving another class of problems. This general class of problems involves attempts to infer the properties of a large collection of data from inspection of a sample of observations. For example, a school superin-

tendent wishes to determine the proportion of children in a large school system who come in without breakfast (have used drugs, have been vaccinated for Asian flu, or whatever). Having a little knowledge of statistics, the superintendent would know that it is unnecessary and inefficient to question each child; the proportion for the entire district could be estimated fairly accurately from a sample of as few as 100 children.

Thus the purpose of inferential statistics is to predict or estimate characteristics of a population from a knowledge of the characteristics of only a sample of the population. The descriptive characteristics of a sample can be generalized to the entire population, with a known margin of error, using the techniques of inferential statistics.

The design and analysis of experiments is an important branch of inferential statistical methods. These methods were developed for testing causal relationships among variables. Experimental design is so important for the study of causal relationships that in some philosophical systems an experiment constitutes an operational definition of a causal relationship. Adults make causal inferences during all their waking moments. The frequent use of the word “because” reveals this: “The school bond failed to pass because it was not well publicized” or “He scored poorly on the intelligence test because he was overly anxious about the consequences of the score.” Statistical methods assist researchers in describing data, in drawing inferences to larger bodies of data, and in studying causal relationships.

1.4

STATISTICS AND MATHEMATICS

The discipline of statistics is a branch of applied mathematics. Mastering statistical methods requires some mathematical proficiency, but less than commonly assumed. Do not think statistics is accessible only to the specially trained. In this book, much use is made of intuition, logical reasoning, and simple arithmetic. Much of the rationale of applied statistics and many of its techniques can be learned without advanced mathematical skills.

If you have not studied mathematics, logic, or any other rigorous and deductive discipline recently, you may find studying statistics uncomfortable for a while. In many disciplines characterized by vague verbal discourse and personalistic use of language, a student can sustain sloppy and erroneous thinking for long periods without being aware of it. A speaker might receive an enthusiastic audience reaction to the statement, “Viable individualized, democratic, and creative alternatives are necessary to meet the needs of the whole child.” If the statement is scrutinized, however, its meaning is so ambiguous and imprecise that it is essentially meaningless. The student of statistics is likely to be confronted abruptly and uncomfortably with the results of careless thinking. If you are inclined toward critical and precise thought, this restrictive and confining mantle will soon begin to feel comfortable. The satisfying reassurance of knowing that you are mastering a logical and unambiguous language will outweigh the work involved in learning it. Being wrong on occasion is the price we must pay for knowing when we are correct. Not knowing if we are speaking nonsense is too expensive a luxury to entertain in an age in which sense is scarce.

A word to the wise: “Be ye doers of the word, and not hearers only, deceiving your own selves” (James 1:22). By far the greatest demand that the study of statistics exacts from the student is thorough, detailed, and careful attention to the subject. A quick reading of this book will not produce a mastery of statistics. A statistics text is not