Statistical Methods for the Social Sciences

Fifth Edition

Alan Agresti

University of Florida

Pearson

Contents

Preface ix

Acknowledgments xi

1 Introduction 1

1.1 Introduction to Statistical Methodology 1

1.2 Descriptive Statistics and Inferential Statistics 4

1.3 The Role of Computers and Software in Statistics 6

1.4 Chapter Summary 8

2 Sampling and Measurement 11

2.1 Variables and Their Measurement 11

2.2 Randomization 14

2.3 Sampling Variability and Potential Bias 17

2.4 Other Probability Sampling Methods* 21

2.5 Chapter Summary 23

3 Descriptive Statistics 29

3.1 Describing Data with Tables and Graphs 29

3.2 Describing the Center of the Data 35

3.3 Describing Variability of the Data 41

3.4 Measures of Position 46

3.5 Bivariate Descriptive Statistics 51

3.6 Sample Statistics and Population Parameters 55

3.7 Chapter Summary 55

4 Probability Distributions 67

4.1 Introduction to Probability 67

4.2 Probability Distributions for Discrete and Continuous Variables 69

4.3 The Normal Probability Distribution 72

4.4 Sampling Distributions Describe How Statistics Vary 80

4.5 Sampling Distributions of Sample Means 85

4.6 Review: Population, Sample Data, and Sampling Distributions 91

4.7 Chapter Summary 94

5 Statistical Inference: Estimation 103

5.1 Point and Interval Estimation 103

5.2 Confidence Interval for a Proportion 106

5.3 Confidence Interval for a Mean 113

5.4 Choice of Sample Size 120

5.5 Estimation Methods: Maximum Likelihood and the Bootstrap* 126

5.6 Chapter Summary 130

6 Statistical Inference: Significance Tests 139

6.1 The Five Parts of a Significance Test 140

6.2 Significance Test for a Mean 143

6.3 Significance Test for a Proportion 152

6.4 Decisions and Types of Errors in Tests 155

6.5 Limitations of Significance Tests 159

6.6 Finding *P*(Type II Error)* 163

6.7 Small-Sample Test for a Proportion—the Binomial Distribution* 165

6.8 Chapter Summary 169

7 Comparison of Two Groups 179

7.1 Preliminaries for Comparing Groups 179

7.2 Categorical Data: Comparing Two Proportions 182

7.3 Quantitative Data: Comparing Two Means 187

7.4 Comparing Means with Dependent Samples 190
7.5 Other Methods for Comparing Means* 193
7.6 Other Methods for Comparing Proportions* 198
7.7 Nonparametric Statistics for Comparing Groups* 201
7.8 Chapter Summary 204

8 Analyzing Association between Categorical Variables 215

8.1 Contingency Tables 215
8.2 Chi-Squared Test of Independence 218
8.3 Residuals: Detecting the Pattern of Association 225
8.4 Measuring Association in Contingency Tables 227
8.5 Association Between Ordinal Variables* 233
8.6 Chapter Summary 238

9 Linear Regression and Correlation 247

9.1 Linear Relationships 247
9.2 Least Squares Prediction Equation 250
9.3 The Linear Regression Model 256
9.4 Measuring Linear Association: The Correlation 259
9.5 Inferences for the Slope and Correlation 266
9.6 Model Assumptions and Violations 272
9.7 Chapter Summary 277

10 Introduction to Multivariate Relationships 287

10.1 Association and Causality 287
10.2 Controlling for Other Variables 290
10.3 Types of Multivariate Relationships 294
10.4 Inferential Issues in Statistical Control 299
10.5 Chapter Summary 301

11 Multiple Regression and Correlation 307

11.1 The Multiple Regression Model 307
11.2 Multiple Correlation and R^2 316
11.3 Inferences for Multiple Regression Coefficients 320
11.4 Modeling Interaction Effects 325
11.5 Comparing Regression Models 329
11.6 Partial Correlation* 331
11.7 Standardized Regression Coefficients* 334
11.8 Chapter Summary 337

12 Regression with Categorical Predictors: Analysis of Variance Methods 351

12.1 Regression Modeling with Dummy Variables for Categories 351
12.2 Multiple Comparisons of Means 355
12.3 Comparing Several Means: Analysis of Variance 358
12.4 Two-Way ANOVA and Regression Modeling 362
12.5 Repeated-Measures Analysis of Variance* 369
12.6 Two-Way ANOVA with Repeated Measures on a Factor* 373
12.7 Chapter Summary 378

13 Multiple Regression with Quantitative and Categorical Predictors 387

13.1 Models with Quantitative and Categorical Explanatory Variables 387
13.2 Inference for Regression with Quantitative and Categorical Predictors 394
13.3 Case Studies: Using Multiple Regression in Research 397

13.4 Adjusted Means* 401
13.5 The Linear Mixed Model* 406
13.6 Chapter Summary 411

14 Model Building with Multiple Regression 419

14.1 Model Selection Procedures 419
14.2 Regression Diagnostics 426
14.3 Effects of Multicollinearity 433
14.4 Generalized Linear Models 435
14.5 Nonlinear Relationships: Polynomial Regression 439
14.6 Exponential Regression and Log Transforms* 444
14.7 Robust Variances and Nonparametric Regression* 448
14.8 Chapter Summary 450

15 Logistic Regression: Modeling Categorical Responses 459

15.1 Logistic Regression 459
15.2 Multiple Logistic Regression 465
15.3 Inference for Logistic Regression Models 470
15.4 Logistic Regression Models for Ordinal Variables* 472
15.5 Logistic Models for Nominal Responses* 477
15.6 Loglinear Models for Categorical Variables* 480
15.7 Model Goodness-of-Fit Tests for Contingency Tables* 484
15.8 Chapter Summary 488

16 An Introduction to Advanced Methodology 497

16.1 Missing Data: Adjustment Using Multiple Imputation* 497
16.2 Multilevel (Hierarchical) Models* 501
16.3 Event History Models* 503
16.4 Path Analysis* 506
16.5 Factor Analysis* 510
16.6 Structural Equation Models* 515
16.7 Markov Chains* 519
16.8 The Bayesian Approach to Statistical Inference* 520

Appendix: R, Stata, SPSS, and SAS for Statistical Analyses 527

Answers to Selected Odd-Numbered Exercises 565

Bibliography 579

Credits 583

Index 585