An Introduction to Mathematical Statistics and Its Applications

Sixth Edition

Richard J. Larsen
Vanderbilt University

Morris L. Marx
University of West Florida

Pearson

Contents

Preface viii

1 Introduction 1

1.1 An Overview 1
1.2 Some Examples 2
1.3 A Brief History 6
1.4 A Chapter Summary 14

2 Probability 15

2.1 Introduction 15
2.2 Sample Spaces and the Algebra of Sets 17
2.3 The Probability Function 26
2.4 Conditional Probability 31
2.5 Independence 50
2.6 Combinatorics 65
2.7 Combinatorial Probability 89
2.8 Taking a Second Look at Statistics (Monte Carlo Techniques) 99

3 Random Variables 102

3.1 Introduction 102
3.2 Binomial and Hypergeometric Probabilities 103
3.3 Discrete Random Variables 116
3.4 Continuous Random Variables 127
3.5 Expected Values 137
3.6 The Variance 153
3.7 Joint Densities 160
3.8 Transforming and Combining Random Variables 174
3.9 Further Properties of the Mean and Variance 182
3.10 Order Statistics 192
3.11 Conditional Densities 199
3.12 Moment-Generating Functions 206
3.13 Taking a Second Look at Statistics (Interpreting Means) 215

4 Special Distributions 218

4.1 Introduction **218**

4.2 The Poisson Distribution **219**

4.3 The Normal Distribution **235**

4.4 The Geometric Distribution **257**

4.5 The Negative Binomial Distribution **259**

4.6 The Gamma Distribution **267**

4.7 Taking a Second Look at Statistics (Monte Carlo Simulations) **271**

Appendix 4.A.1 Properties of Frequently-Used pdfs **274**

Appendix 4.A.2 A Proof of the Central Limit Theorem **276**

5 Estimation 278

5.1 Introduction **278**

5.2 Estimating Parameters: The Method of Maximum Likelihood and the Method of Moments **280**

5.3 Interval Estimation **293**

5.4 Properties of Estimators **308**

5.5 Minimum-Variance Estimators: The Cramér-Rao Lower Bound **316**

5.6 Sufficient Estimators **319**

5.7 Consistency **326**

5.8 Bayesian Estimation **329**

5.9 Taking a Second Look at Statistics (Beyond Classical Estimation) **341**

6 Hypothesis Testing 343

6.1 Introduction **343**

6.2 The Decision Rule **344**

6.3 Testing Binomial Data—H_0: $p = p_0$ **353**

6.4 Type I and Type II Errors **359**

6.5 A Notion of Optimality: The Generalized Likelihood Ratio **375**

6.6 Taking a Second Look at Hypothesis Testing (Statistical Significance versus "Practical" Significance) **378**

7 Inferences Based on the Normal Distribution 380

7.1 Introduction **380**

7.2 Comparing $\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}}$ and $\frac{\bar{Y}-\mu}{S/\sqrt{n}}$ **381**

7.3 Deriving the Distribution of $\frac{\bar{Y}-\mu}{S/\sqrt{n}}$ **383**

7.4 Drawing Inferences About μ **389**

7.5 Drawing Inferences About σ^2 **404**

7.6 Taking a Second Look at Statistics (Type II Error) **412**

Appendix 7.A.1 Some Distribution Results for $\bar{Y}$ and S^2 **414**

Appendix 7.A.2 A Proof That the One-Sample t Test Is a GLRT **416**

Appendix 7.A.3 A Proof of Theorem 7.5.2 **418**

8 Types of Data: A Brief Overview **421**

8.1 Introduction **421**

8.2 Classifying Data **427**

8.3 Taking a Second Look at Statistics (Why Samples Are Not "Valid"!) **448**

9 Two-Sample Inferences **450**

9.1 Introduction **450**

9.2 Testing H_0: $\mu_X = \mu_Y$ **451**

9.3 Testing H_0: $\sigma_X^2 = \sigma_Y^2$—The F Test **463**

9.4 Binomial Data: Testing H_0: $p_X = p_Y$ **468**

9.5 Confidence Intervals for the Two-Sample Problem **473**

9.6 Taking a Second Look at Statistics (Choosing Samples) **478**

Appendix 9.A.1 A Derivation of the Two-Sample t Test (A Proof of Theorem 9.2.2) **480**

10 Goodness-of-Fit Tests **483**

10.1 Introduction **483**

10.2 The Multinomial Distribution **484**

10.3 Goodness-of-Fit Tests: All Parameters Known **488**

10.4 Goodness-of-Fit Tests: Parameters Unknown **498**

10.5 Contingency Tables **507**

10.6 Taking a Second Look at Statistics (Outliers) **517**

11 Regression **520**

11.1 Introduction **520**

11.2 The Method of Least Squares **520**

11.3 The Linear Model **543**

11.4 Covariance and Correlation **563**

11.5 The Bivariate Normal Distribution **570**

11.6 Taking a Second Look at Statistics (How *Not* to Interpret the Sample Correlation Coefficient) **576**

Appendix 11.A.1 A Proof of Theorem 11.3.3 **577**

12 THE ANALYSIS OF VARIANCE **580**

12.1 Introduction **580**

12.2 The *F* Test **582**

12.3 Multiple Comparisons: Tukey's Method **592**

12.4 Testing Subhypotheses with Contrasts **596**

12.5 Data Transformations **604**

12.6 Taking a Second Look at Statistics (Putting the Subject of Statistics Together—The Contributions of Ronald A. Fisher) **606**

Appendix 12.A.1 A Proof of Theorem 12.2.2 **608**

Appendix 12.A.2 The Distribution of $\frac{SSTR/(k-1)}{SSE/(n-k)}$ When H_1 Is True **608**

13 RANDOMIZED BLOCK DESIGNS **613**

13.1 Introduction **613**

13.2 The *F* Test for a Randomized Block Design **614**

13.3 The Paired *t* Test **628**

13.4 Taking a Second Look at Statistics (Choosing between a Two-Sample *t* Test and a Paired *t* Test) **634**

14 NONPARAMETRIC STATISTICS **638**

14.1 Introduction **638**

14.2 The Sign Test **639**

14.3 Wilcoxon Tests **645**

14.4 The Kruskal-Wallis Test **658**

14.5 The Friedman Test **662**

14.6 Testing for Randomness **665**

14.7 Taking a Second Look at Statistics (Comparing Parametric and Nonparametric Procedures) **669**

15 Factorial Data (Available Online) 15-1

15.1 Introduction 15-1
15.2 The Two-Factor Factorial 15-4
15.3 Sums of Squares for Two-Factor Factorials 15-16
15.4 Expected Mean Squares 15-26
15.5 Examples 15-30
15.6 The Three-Factor Factorial Design 15-40
15.7 2^n Designs 15-51
15.8 Fractional Factorials 15-72

Appendix A: Statistical Tables 674

Answers to Selected Odd-Numbered Questions 701

Bibliography 725

Index 737