

Contents at a Glance

<i>Introduction</i>	1
<i>Part I: Tackling Data Analysis and Model-Building Basics</i>	7
Chapter 1: Beyond Number Crunching: The Art and Science of Data Analysis	9
Chapter 2: Finding the Right Analysis for the Job	21
Chapter 3: Reviewing Confidence Intervals and Hypothesis Tests	37
<i>Part II: Using Different Types of Regression to Make Predictions</i>	53
Chapter 4: Getting in Line with Simple Linear Regression	55
Chapter 5: Multiple Regression with Two X Variables	83
Chapter 6: How Can I Miss You If You Won't Leave? Regression Model Selection	103
Chapter 7: Getting Ahead of the Learning Curve with Nonlinear Regression	115
Chapter 8: Yes, No, Maybe So: Making Predictions by Using Logistic Regression	137
<i>Part III: Analyzing Variance with ANOVA</i>	151
Chapter 9: Testing Lots of Means? Come On Over to ANOVA!	153
Chapter 10: Sorting Out the Means with Multiple Comparisons	173
Chapter 11: Finding Your Way through Two-Way ANOVA	191
Chapter 12: Regression and ANOVA: Surprise Relatives!	207
<i>Part IV: Building Strong Connections with Chi-Square Tests</i>	219
Chapter 13: Forming Associations with Two-Way Tables	221
Chapter 14: Being Independent Enough for the Chi-Square Test	241
Chapter 15: Using Chi-Square Tests for Goodness-of-Fit (Your Data, Not Your Jeans)	263
<i>Part V: Nonparametric Statistics: Rebels without a Distribution</i>	273
Chapter 16: Going Nonparametric	275
Chapter 17: All Signs Point to the Sign Test and Signed Rank Test	287

Chapter 18: Pulling Rank with the Rank Sum Test.....	303
Chapter 19: Do the Kruskal-Wallis and Rank the Sums with the Wilcoxon.....	313
Chapter 20: Pointing Out Correlations with Spearman's Rank.....	325
<i>Part VI: The Part of Tens</i>	333
Chapter 21: Ten Common Errors in Statistical Conclusions.....	335
Chapter 22: Ten Ways to Get Ahead by Knowing Statistics.....	347
Chapter 23: Ten Cool Jobs That Use Statistics.....	357
<i>Appendix: Reference Tables</i>	367
<i>Index</i>	379

Table of Contents

.....

Introduction 1

About This Book	1
Conventions Used in This Book.....	2
What You're Not to Read.....	3
Foolish Assumptions.....	3
How This Book Is Organized	3
Part I: Tackling Data Analysis and Model-Building Basics.....	4
Part II: Using Different Types of Regression to Make Predictions...	4
Part III: Analyzing Variance with ANOVA.....	4
Part IV: Building Strong Connections with Chi-Square Tests.....	5
Part V: Nonparametric Statistics: Rebels without a Distribution ...	5
Part VI: The Part of Tens.....	5
Icons Used in This Book	5
Where to Go from Here.....	6

Part I: Tackling Data Analysis and Model-Building Basics 7

Chapter 1: Beyond Number Crunching: The Art and Science of Data Analysis 9

Data Analysis: Looking before You Crunch.....	9
Nothing (not even a straight line) lasts forever.....	11
Data snooping isn't cool	11
No (data) fishing allowed.....	12
Getting the Big Picture: An Overview of Stats II	13
Population parameter	13
Sample statistic	14
Confidence interval.....	14
Hypothesis test	15
Analysis of variance (ANOVA)	15
Multiple comparisons.....	16
Interaction effects.....	16
Correlation.....	17
Linear regression	18
Chi-square tests	19
Nonparametrics	20

Chapter 2: Finding the Right Analysis for the Job 21

Categorical versus Quantitative Variables	22
Statistics for Categorical Variables	23
Estimating a proportion	23
Comparing proportions	24
Looking for relationships between categorical variables.....	25
Building models to make predictions	26
Statistics for Quantitative Variables	27
Making estimates	27
Making comparisons	28
Exploring relationships.....	28
Predicting y using x	30
Avoiding Bias	31
Measuring Precision with Margin of Error	33
Knowing Your Limitations.....	34

Chapter 3: Reviewing Confidence Intervals and Hypothesis Tests. 37

Estimating Parameters by Using Confidence Intervals.....	38
Getting the basics: The general form of a confidence interval	38
Finding the confidence interval for a population mean.....	39
What changes the margin of error?.....	40
Interpreting a confidence interval	43
What's the Hype about Hypothesis Tests?	44
What H_0 and H_a really represent	44
Gathering your evidence into a test statistic	45
Determining strength of evidence with a p-value.....	45
False alarms and missed opportunities: Type I and II errors.....	46
The power of a hypothesis test	48

Part II: Using Different Types of Regression to Make Predictions*..... 53*Chapter 4: Getting in Line with Simple Linear Regression. 55**

Exploring Relationships with Scatterplots and Correlations.....	56
Using scatterplots to explore relationships	57
Collating the information by using the correlation coefficient.....	58
Building a Simple Linear Regression Model.....	60
Finding the best-fitting line to model your data.....	60
The y-intercept of the regression line.....	61
The slope of the regression line.....	62
Making point estimates by using the regression line.....	63

No Conclusion Left Behind: Tests and Confidence Intervals
 for Regression 63
 Scrutinizing the slope..... 64
 Inspecting the y-intercept..... 66
 Building confidence intervals for the average response 68
 Making the band with prediction intervals 69
 Checking the Model’s Fit (The Data, Not the Clothes!) 71
 Defining the conditions..... 71
 Finding and exploring the residuals 73
 Using r^2 to measure model fit..... 76
 Scoping for outliers 77
 Knowing the Limitations of Your Regression Analysis..... 79
 Avoiding slipping into cause-and-effect mode 79
 Extrapolation: The ultimate no-no..... 80
 Sometimes you need more than one variable..... 81

Chapter 5: Multiple Regression with Two X Variables. 83

Getting to Know the Multiple Regression Model..... 83
 Discovering the uses of multiple regression 84
 Looking at the general form of the multiple regression model..... 84
 Stepping through the analysis 85
 Looking at x’s and y’s..... 85
 Collecting the Data 86
 Pinpointing Possible Relationships..... 88
 Making scatterplots 88
 Correlations: Examining the bond..... 89
 Checking for Multicollinearity 91
 Finding the Best-Fitting Model for Two x Variables 92
 Getting the multiple regression coefficients 93
 Interpreting the coefficients 94
 Testing the coefficients..... 95
 Predicting y by Using the x Variables 97
 Checking the Fit of the Multiple Regression Model 98
 Noting the conditions..... 98
 Plotting a plan to check the conditions 98
 Checking the three conditions..... 100

**Chapter 6: How Can I Miss You If You Won’t Leave?
 Regression Model Selection 103**

Getting a Kick out of Estimating Punt Distance 104
 Brainstorming variables and collecting data 104
 Examining scatterplots and correlations..... 106



Just Like Buying Shoes: The Model Looks Nice, But Does It Fit? 109
 Assessing the fit of multiple regression models 110
 Model selection procedures 111

**Chapter 7: Getting Ahead of the Learning Curve
with Nonlinear Regression 115**

Anticipating Nonlinear Regression 116
Starting Out with Scatterplots 117
Handling Curves in the Road with Polynomials 119
 Bringing back polynomials 119
 Searching for the best polynomial model 122
 Using a second-degree polynomial to pass the quiz 123
 Assessing the fit of a polynomial model 126
 Making predictions 129
Going Up? Going Down? Go Exponential! 130
 Recollecting exponential models 130
 Searching for the best exponential model 131
 Spreading secrets at an exponential rate 133

**Chapter 8: Yes, No, Maybe So: Making Predictions
by Using Logistic Regression 137**

Understanding a Logistic Regression Model 138
 How is logistic regression different from other regressions? 138
 Using an S-curve to estimate probabilities 139
 Interpreting the coefficients of the logistic regression model 140
 The logistic regression model in action 141
Carrying Out a Logistic Regression Analysis 142
 Running the analysis in Minitab 142
 Finding the coefficients and making the model 144
 Estimating p 145
 Checking the fit of the model 146
 Fitting the Movie Model 147

Part III: Analyzing Variance with ANOVA 151

Chapter 9: Testing Lots of Means? Come On Over to ANOVA! 153

Comparing Two Means with a t-Test 154
Evaluating More Means with ANOVA 155
 Spitting seeds: A situation just waiting for ANOVA 155
 Walking through the steps of ANOVA 156
Checking the Conditions 157
 Verifying independence 157
 Looking for what's normal 158
 Taking note of spread 159
Setting Up the Hypotheses 162

Doing the F-Test..... 162
 Running ANOVA in Minitab 163
 Breaking down the variance into sums of squares..... 164
 Locating those mean sums of squares..... 165
 Figuring the F-statistic..... 166
 Making conclusions from ANOVA..... 168
 What’s next? 169
 Checking the Fit of the ANOVA Model..... 170

Chapter 10: Sorting Out the Means with Multiple Comparisons. . . . 173

Following Up after ANOVA 174
 Comparing cellphone minutes: An example..... 174
 Setting the stage for multiple comparison procedures 176
 Pinpointing Differing Means with Fisher and Tukey 177
 Fishing for differences with Fisher’s LSD..... 178
 Using Fisher’s new and improved LSD..... 179
 Separating the turkeys with Tukey’s test 182
 Examining the Output to Determine the Analysis 183
 So Many Other Procedures, So Little Time! 184
 Controlling for baloney with the Bonferroni adjustment 185
 Comparing combinations by using Scheffe’s method..... 186
 Finding out whodunit with Dunnett’s test..... 186
 Staying cool with Student Newman-Keuls 187
 Duncan’s multiple range test..... 187
 Going nonparametric with the Kruskal-Wallis test..... 188

Chapter 11: Finding Your Way through Two-Way ANOVA 191

Setting Up the Two-Way ANOVA Model 192
 Determining the treatments 192
 Stepping through the sums of squares 193
 Understanding Interaction Effects..... 194
 What is interaction, anyway? 195
 Interacting with interaction plots..... 195
 Testing the Terms in Two-Way ANOVA..... 198
 Running the Two-Way ANOVA Table 199
 Interpreting the results: Numbers and graphs 200
 Are Whites Whiter in Hot Water? Two-Way ANOVA Investigates 202

Chapter 12: Regression and ANOVA: Surprise Relatives! 207

Seeing Regression through the Eyes of Variation 208
 Spotting variability and finding an “x-planation” 208
 Getting results with regression..... 209
 Assessing the fit of the regression model..... 211
 Regression and ANOVA: A Meeting of the Models 212
 Comparing sums of squares 212
 Dividing up the degrees of freedom 214
 Bringing regression to the ANOVA table 215
 Relating the F- and t-statistics: The final frontier 216

***Part IV: Building Strong Connections with Chi-Square Tests* 219**

Chapter 13: Forming Associations with Two-Way Tables 221

Breaking Down a Two-Way Table	222
Organizing data into a two-way table	222
Filling in the cell counts	223
Making marginal totals	224
Breaking Down the Probabilities	225
Marginal probabilities	226
Joint probabilities	227
Conditional probabilities	228
Trying To Be Independent	233
Checking for independence between two categories	233
Checking for independence between two variables	235
Demystifying Simpson's Paradox	236
Experiencing Simpson's Paradox	236
Figuring out why Simpson's Paradox occurs	239
Keeping one eye open for Simpson's Paradox	240

Chapter 14: Being Independent Enough for the Chi-Square Test . . 241

The Chi-square Test for Independence	242
Collecting and organizing the data	243
Determining the hypotheses	245
Figuring expected cell counts	245
Checking the conditions for the test	246
Calculating the Chi-square test statistic	247
Finding your results on the Chi-square table	249
Drawing your conclusions	253
Putting the Chi-square to the test	255
Comparing Two Tests for Comparing Two Proportions	257
Getting reacquainted with the Z-test for two population proportions	257
Equating Chi-square tests and Z-tests for a two-by-two table	258

Chapter 15: Using Chi-Square Tests for Goodness-of-Fit (Your Data, Not Your Jeans) 263

Finding the Goodness-of-Fit Statistic	264
What's observed versus what's expected	264
Calculating the goodness-of-fit statistic	266
Interpreting the Goodness-of-Fit Statistic Using a Chi-Square	268
Checking the conditions before you start	270
The steps of the Chi-square goodness-of-fit test	270

***Part V: Nonparametric Statistics:
Rebels without a Distribution* 273**

Chapter 16: Going Nonparametric275

Arguing for Nonparametric Statistics	275
No need to fret if conditions aren't met.....	276
The median's in the spotlight for a change.....	277
So, what's the catch?.....	279
Mastering the Basics of Nonparametric Statistics	280
Sign	280
Rank.....	282
Signed rank	283
Rank sum.....	284

**Chapter 17: All Signs Point to the Sign Test
and Signed Rank Test287**

Reading the Signs: The Sign Test.....	288
Testing the median.....	290
Estimating the median.....	292
Testing matched pairs.....	294
Going a Step Further with the Signed Rank Test.....	296
A limitation of the sign test	296
Stepping through the signed rank test.....	297
Losing weight with signed ranks.....	298

Chapter 18: Pulling Rank with the Rank Sum Test. 303

Conducting the Rank Sum Test.....	303
Checking the conditions	303
Stepping through the test	304
Stepping up the sample size.....	306
Performing a Rank Sum Test: Which Real Estate Agent Sells Homes Faster?	307
Checking the conditions for this test	307
Testing the hypotheses.....	309

**Chapter 19: Do the Kruskal-Wallis and
Rank the Sums with the Wilcoxon313**

Doing the Kruskal-Wallis Test to Compare More than Two Populations	313
Checking the conditions	315
Setting up the test.....	317
Conducting the test step by step.....	317

Pinpointing the Differences: The Wilcoxon Rank Sum Test.....	320
Pairing off with pairwise comparisons.....	320
Carrying out comparison tests to see who's different.....	321
Examining the medians to see how they're different.....	323
Chapter 20: Pointing Out Correlations with Spearman's Rank	325
Pickin' On Pearson and His Precious Conditions.....	326
Scoring with Spearman's Rank Correlation.....	327
Figuring Spearman's rank correlation.....	328
Watching Spearman at work: Relating aptitude to performance.....	329
 <i>Part VI: The Part of Tens</i>.....	 333
Chapter 21: Ten Common Errors in Statistical Conclusions	335
Chapter 22: Ten Ways to Get Ahead by Knowing Statistics	347
Chapter 23: Ten Cool Jobs That Use Statistics	357
 <i>Appendix: Reference Tables</i>.....	 367
 <i>Index</i>.....	 379