

BRIEF CONTENTS

CHAPTER ONE

Overview of Statistics 2

CHAPTER TWO

Data Collection 22

CHAPTER THREE

Describing Data Visually 58

CHAPTER FOUR

Descriptive Statistics 112

CHAPTER FIVE

Probability 168

CHAPTER SIX

Discrete Distributions 208

CHAPTER SEVEN

Continuous Distributions 252

CHAPTER EIGHT

Sampling Distributions and Estimation 292

CHAPTER NINE

One-Sample Hypothesis Tests 346

CHAPTER TEN

Two-Sample Hypothesis Tests 394

CHAPTER ELEVEN

Analysis of Variance 438

CHAPTER TWELVE

Bivariate Regression 488

CHAPTER THIRTEEN

Multiple Regression 558

CHAPTER FOURTEEN

Time-Series Analysis 604

CHAPTER FIFTEEN

Chi-Square Tests 656

CHAPTER SIXTEEN

Nonparametric Tests 698

CHAPTER SEVENTEEN

Quality Management 730

CHAPTER EIGHTEEN

Simulation (On Student CD-ROM)

APPENDIXES

A Exact Binomial Probabilities 774

B Exact Poisson Probabilities 776

C-1 Standard Normal Areas 779

C-2 Cumulative Standard Normal
Distribution 780

D Student's t Critical Values 782

E Chi-Square Critical Values 783

F Critical Values of F 784

G Solutions to Odd-Numbered Exercises 792

PHOTO CREDITS 816

INDEX 817

CONTENTS

CHAPTER ONE

Overview of Statistics 2

- 1.1 What is Statistics? 3
- 1.2 Why Study Statistics? 4
 - Communication 4
 - Computer Skills 4
 - Information Management 4
 - Technical Literacy 4
 - Career Advancement 4
 - Quality Improvement 4
- 1.3 Uses of Statistics 5
 - Auditing 5
 - Marketing 5
 - Health Care 5
 - Quality Control 5
 - Purchasing 6
 - Medicine 6
 - Forecasting 6
 - Product Warranty 6
- 1.4 Statistical Challenges 6
 - Working with Imperfect Data 6
 - Dealing with Practical Constraints 7
 - Upholding Ethical Standards 7
 - Using Consultants 7
- 1.5 Writing and Presenting Reports 9
 - Rules for "Power" Writing 9
 - Writing Style 9
 - Spelling and Grammar 10
 - Organizing a Technical Report 10
 - Writing an Executive Summary 10
 - Tables and Graphs 11
 - Rules for Presenting Oral Reports 11
 - The Three Ps 12
- 1.6 Statistical Pitfalls 14
 - Pitfall 1: Making Conclusions about a Large Population from a Small Sample 14
 - Pitfall 2: Making Conclusions from Nonrandom Samples 14
 - Pitfall 3: Attaching Importance to Rare Observations from Large Samples 14
 - Pitfall 4: Using Poor Survey Methods 14
 - Pitfall 5: Assuming a Causal Link Based Only on Observed Association 15
 - Pitfall 6: Making Generalizations about Individuals from Observations about Groups 15
 - Pitfall 7: Unconscious Bias 15
 - Pitfall 8: Attaching Practical Importance to Every Statistically Significant Study Result 15

- 1.7 Statistics: An Evolving Field 16
- Chapter Summary 16

CHAPTER TWO

Data Collection 22

- 2.1 Definitions 23
 - Subjects, Variables, and Data Sets 23
 - Data Types 24
- 2.2 Level of Measurement 26
 - Nominal Measurement 26
 - Ordinal Measurement 27
 - Interval Measurement 27
 - Ratio Measurement 28
 - Changing Data by Recoding 29
- 2.3 Time-Series versus Cross-Sectional Data 30
 - Time-Series Data 30
 - Cross-Sectional Data 30
- 2.4 Sampling Concepts 31
 - Sample or Census? 31
 - Parameters and Statistics 32
 - Target Population 32
 - Finite or Infinite? 33
- 2.5 Sampling Methods 33
 - Simple Random Sample 33
 - Random Number Tables 35
 - Setting Up a Rule 35
 - With or Without Replacement? 35
 - Computer Methods 36
 - Row/Column Data Arrays 36
 - Randomizing a List 37
 - Systematic Sample 37
 - Stratified Sample 39
 - Applications of Stratified Sampling 39
 - Cluster Sample 39
 - Judgment Sample 40
 - Convenience Sample 40
 - Sample Size 41
- 2.6 Data Sources 42
- 2.7 Survey Research 43
 - Survey Types 43
 - Response Rates 43
 - Getting Advice 44
 - Questionnaire Design 44
 - Question Wording 44
 - Coding and Data Screening 46
 - Sources of Error 46
 - Data File Format 47
 - Chapter Summary 49

CHAPTER THREE

Describing Data Visually 58

3.1 Visual Description 59

Measurement 60

Sorting 60

3.2 Dot Plots 61

Small Sample: Home Prices 61

Comparing Groups 63

3.3 Frequency Distributions and Histograms 65

Bins and Bin Limits 65

Constructing a Frequency Distribution 65

Histograms 66

Excel Histograms 66

MegaStat Histograms 68

MINITAB Histograms 68

Modal Class 69

Shape 69

3.4 Line Charts 72

Simple Line Charts 72

Grid Lines 73

Log Scales 73

Tips for Effective Line Charts 75

3.5 Bar Charts 76

Plain Bar Charts 76

3-D and Novelty Bar Charts 76

Pareto Charts 77

Stacked Bar Chart 78

Bar Charts for Time-Series Data 78

Tips for Effective Bar Charts 79

3.6 Scatter Plots 80

Policy Making 82

Degree of Association 82

Making a Scatter Plot in Excel 84

3.7 Tables 86

Tips for Effective Tables 87

3.8 Pie Charts 87

An Oft-Abused Chart 87

Pie Chart Options 87

3.9 Effective Excel Charts 90

Chart Wizard 90

Embellished Charts 91

3.10 Maps and Pictograms 94

Spatial Variation and GIS 94

Pictograms 95

3.11 Deceptive Graphs 95

Error 1: Nonzero Origin 95

Error 2: Elastic Graph Proportions 96

Error 3: Dramatic Title 96

Error 4: Distracting Pictures 96

Error 5: Authority Figures 97

Error 6: 3-D and Rotated Graphs 97

Error 7: Missing Axis Demarcations 97

Error 8: Missing Measurement Units

or Definitions 97

Error 9: Vague Source 97

Error 10: Complex Graphs 97

Error 11: Gratuitous Effects 98

Error 12: Estimated Data 98

Error 13: Area Trick 98

Final Advice 98

Further Challenges 99

Chapter Summary 99

CHAPTER FOUR

Descriptive Statistics 112

4.1 Numerical Description 113

Preliminary Analysis 114

Sorting 114

Visual Displays 114

Descriptive Statistics in Excel 116

Descriptive Statistics in MegaStat 117

4.2 Central Tendency 119

Mean 119

Characteristics of the Mean 119

Median 120

Characteristics of the Median 121

Mode 121

Skewness 123

Geometric Mean 127

Growth Rates 127

Midrange 128

Trimmed Mean 128

4.3 Dispersion 130

Range 131

Variance 131

Standard Deviation 131

Calculating a Standard Deviation 132

Characteristics of the Standard

Deviation 133

Coefficient of Variation 133

Mean Absolute Deviation 133

Central Tendency versus Dispersion:

Manufacturing 135

Central Tendency and Dispersion:

Job Performance 135

4.4 Standardized Data 136

Chebyshev's Theorem 136

The Empirical Rule 137

Unusual Observations 137

Defining a Standardized Variable 138

Outliers 139

Estimating Sigma 139

4.5 Percentiles and Quartiles 141

Percentiles 141

Quartiles 141

Method of Medians 142

Formula Method 143

Excel Quartiles 143

Dispersion Using Quartiles 144

Midhinge 144

Midsread (Interquartile Range) 145

Coefficient of Quartile Variation 145

4.6 Box Plots 145

Fences and Unusual Data Values 146

4.7 Grouped Data 147

Nature and Grouped Data 147

Mean and Standard Deviation 148

Accuracy Issues 148

Properties of Grouped Estimates 149

4.8 Skewness and Kurtosis 149

Skewness 149

Kurtosis 150

Chapter Summary 152

CHAPTER FIVE

Probability 168

5.1 Random Experiments 169

Sample Space 169

Events 170

5.2 Probability 171

Definitions 171

What Is "Probability"? 171

Empirical Approach 172

Law of Large Numbers 172

Practical Issues for Actuaries 172

Classical Approach 173

Subjective Approach 174

5.3 Rules of Probability 174

Complement of an Event 174

Odds of an Event 175

Union of Two Events 175

Intersection of Two Events 175

General Law of Addition 176

Mutually Exclusive Events 177

Special Law of Addition 177

Collectively Exhaustive Sets 177

Forced Dichotomy 177

Conditional Probability 178

5.4 Independent Events 180

Dependent Events 180

Actuaries Again 181

Multiplication Law for Independent Events 181

The Five Nines Rule 181

How Much Redundancy Is Needed? 182

Applications of Redundancy 182

5.5 Contingency Tables 183

What Is a Contingency Table? 183

Marginal Probabilities 184

Joint Probabilities 184

Conditional Probabilities 185

Independence 185

Relative Frequencies 186

How Do We Get a Contingency Table? 187

5.6 Tree Diagrams 190

What Is a Tree? 190

5.7 Bayes's Theorem (Optional) 191

How Bayes's Theorem Works 191

General Form of Bayes's Theorem 192

5.8 Counting Rules (Optional) 196

Fundamental Rule of Counting 196

Factorials 197

Permutations 198

Combinations 198

Chapter Summary 200

CHAPTER SIX

Discrete Distributions 208

6.1 Probability Models 209

6.2 Discrete Distributions 209

Random Variables 209

Probability Distributions 210

Expected Value 211

Application: Life Insurance 212

Application: Raffle Tickets 212

Actuarial Fairness 213

Variance and Standard Deviation 213

What Is a PDF or CDF? 214

6.3 Uniform Distribution 215

Characteristics of the Uniform Distribution 215

Application: Pumping Gas 217

Uniform Random Integers 218

Application: Copier Codes 218

Uniform Model in LearningStats 219

6.4 Bernoulli Distribution 219

Bernoulli Experiments 219

6.5 Binomial Distribution 220

Characteristics of the Binomial Distribution 220

Binomial Shape 221

Application: Uninsured Patients 222

Using the Binomial Formula 222

Using Tables: Appendix A 224

Compound Events 224

Binomial Probabilities: Excel 225

Binomial Probabilities: MegaStat 225

Binomial Probabilities: Visual Statistics 225

Binomial Probabilities: LearningStats 225

Binomial Random Data 225

Recognizing Binomial Applications 225

6.6 Poisson Distribution 228

Poisson Processes 228

Characteristics of the Poisson Distribution 229

Using the Poisson Formula 230

Compound Events 232

Poisson Probabilities: Tables (Appendix B) 232

Poisson Probabilities: Excel 232

Poisson Probabilities: Visual Statistics 232

Recognizing Poisson Applications 232

Poisson Approximation to Binomial

(Optional) 234



6.7	Hypergeometric Distribution	235
	Characteristics of the Hypergeometric Distribution	235
	Using the Hypergeometric Formula	236
	Hypergeometric Probabilities: Excel	237
	Hypergeometric Probabilities: Visual Statistics	237
	Hypergeometric Probabilities: LearningStats	238
	Recognizing Hypergeometric Applications	238
	Binomial Application to the Hypergeometric (Optional)	239
6.8	Geometric Distribution (Optional)	240
	Characteristics of the Geometric Distribution	240
	Using LearningStats	241
6.9	Transformations of Random Variables (Optional)	242
	Linear Transformation	242
	Application: Exam Scores	242
	Application: Total Cost	242
	Sums of Random Variables	243
	Application: Gasoline Expenses	243
	Application: Project Scheduling	243
	Chapter Summary	244

CHAPTER SEVEN

Continuous Distributions 252

7.1	Continuous Variables	253
	Events as Intervals	253
7.2	Describing a Continuous Distribution	253
	PDFs and CDFs	253
	Probabilities as Areas	254
	Expected Value and Variance	255
	Oh My, Calculus?	255
7.3	Uniform Continuous Distribution	255
	Characteristics of the Uniform Distribution	255
	Special Case: Unit Rectangular	257
	Uses of the Uniform Model	258
7.4	Normal Distribution	258
	Characteristics of the Normal Distribution	258
	What Is Normal?	260
7.5	Standard Normal Distribution	261
	Characteristics of the Standard Normal	261
	Normal Areas from Appendix C-1	262
	Basis for the Empirical Rule	263
	Normal Areas from Appendix C-2	264
	Finding z for a Given Area	265
	Finding Normal Areas with Excel	267
	Finding Areas by Using Standardized Variables	267
	Inverse Normal	269
	Using Excel Without Standardizing	270
	Normal Random Data (Optional)	271
7.6	Normal Approximation to the Binomial (Optional)	273
	When Is Approximation Needed?	273

7.7	Normal Approximation to the Poisson (Optional)	276
	When Is Approximation Needed?	276
7.8	Exponential Distribution	277
	Characteristics of the Exponential Distribution	277
	Inverse Exponential	279
	Mean Time Between Events	280
	Using Excel	281
7.9	Triangular Distribution (Optional)	282
	Characteristics of the Triangular Distribution	282
	Special Case: Symmetric Triangular	284
	Uses of the Triangular	284
	Chapter Summary	284

CHAPTER EIGHT

Sampling Distributions and Estimation 292

8.1	Sampling Variation	293
8.2	Estimators and Sampling Distributions	295
	Some Terminology	295
	Sampling Distributions	295
	Bias	295
	Efficiency	297
	Consistency	297
8.3	Sample Mean and the Central Limit Theorem	298
	Central Limit Theorem for a Mean	299
	Symmetric Population: Uniform Distribution	299
	Skewed Population: Waiting Time	300
	Range of Sample Means	302
	Illustration: GMAT Scores	303
	Sample Size and Standard Error	304
	Illustration: All Possible Samples from a Uniform Population	304
8.4	Confidence Interval for a Mean (μ) with Known σ	306
	What Is a Confidence Interval?	306
	Choosing a Confidence Interval	308
	Is σ Ever Known?	308
8.5	Confidence Interval for a Mean (μ) with Unknown σ	309
	Student's t Distribution	309
	Degrees of Freedom	310
	Comparison of z and t	310
	Confidence Interval Width	313
	A "Good" Sample?	313
	More Analysis Needed	314
	Messy Data?	314
	Using Appendix D	315
	Using Excel	315
	Using MegaStat	316
	Using MINITAB	316
8.6	Confidence Interval for a Proportion (π)	317
	Illustration: Internet Hotel Reservations	318
	Applying the CLT	318
	When Is It Safe to Assume Normality?	320
	Standard Error of the Proportion	320

Confidence Interval for π	321
Narrowing the Interval?	322
Using Excel and MegaStat	323
Small Samples: MINITAB	323
Polls and Margin of Error	324
Rule of Three	324
Very Quick Rule	325
Advice on Proportions	325
8.7 Sample Size Determination for a Mean	326
A Myth	326
Sample Size to Estimate μ	326
How to Estimate σ	327
Using LearningStats	328
Using MegaStat	328
Caution 1: Units of Measure	328
Caution 2: Using z	328
Caution 3: Larger n Is Better	328
8.8 Sample Size Determination for a Proportion	329
Alternatives	330
Practical Advice	330
Using LearningStats	331
Caution 1: Units of Measure	331
Caution 2: Finite Population	331
8.9 Confidence Interval for the Difference of Two Means, $\mu_1 - \mu_2$ (Optional)	331
Should Sample Sizes Be Equal?	333
8.10 Confidence Interval for the Difference of Two Proportions, $\pi_1 - \pi_2$ (Optional)	334
8.11 Confidence Interval for a Population Variance, σ^2 (Optional)	335
Chi-Square Distribution	335
Confidence Interval for σ	336
Using LearningStats	336
Caution: Assumption of Normality	336
Chapter Summary	337

CHAPTER NINE

One-Sample Hypothesis Tests 346

9.1 Logic of Hypothesis Testing	347
Process of Science	347
Who Tests Hypotheses?	348
Good News	348
Hypothesis Formulation	348
Can Hypotheses Be Proved?	348
Role of Evidence	349
Types of Error	349
Statistical Hypothesis Testing	350
One-Sided Tests	351
When to Use a One-Sided Test	352
Decision Rule	352
Type I Error	352
Type II Error	354
Power of a Test	354
Relationship Between α and β	355
Consequences of Type II Error	355
Choice of α	355

Statistical Significance versus Practical Importance 355

9.2 Testing a Proportion	357
Critical Value	358
p-Value Method	359
Interpreting a p-Value	360
Two-Tailed Test	360
Calculating a p-Value for a Two-Tailed Test	361
Effect of α	362
Using the p-Value	363
Effect of a Larger Sample	364
Small Samples and Non-Normality (Optional)	366
9.3 Testing a Mean: Known Population Variance	367
Test Statistic	367
One-Tailed Test	368
p-Value Method	369
Two-Tailed Test	369
Using the p-Value	370
Analogy to Confidence Intervals	371
Significance versus Importance	371
9.4 Testing a Mean: Unknown Population Variance	372
Using Student's t	372
Sensitivity to α	373
Using the p-Value	373
Significance versus Importance	374
Normality Assumption	374
Confidence Interval versus Hypothesis Test	374
Using MegaStat	375
Large Samples	375
9.5 Power Curves and OC Curves (Optional)	377
Power Curve for a Mean: An Example	377
Calculating Power	378
Effect of Sample Size	380
Relationship of the Power and OC Curves	380
Power Curve for Tests of a Proportion	381
Using LearningStats	383
Using Visual Statistics	383
9.6 Tests for One Variance (Optional)	384
Using MegaStat	385
When to Use Tests for One Variance	386
Chapter Summary	387

CHAPTER TEN

Two-Sample Hypothesis Tests 394

10.1 Two-Sample Tests	395
What Is a Two Sample Test?	395
Basis of Two-Sample Tests	396
Test Procedure	396
10.2 Comparing Two Proportions	397
Testing for Zero Difference: $\pi_1 = \pi_2$	397
Sample Proportions	397
Pooled Proportion	397
Test Statistic	397
Using the p-Value	399
Checking Normality	399

Small Samples 400
Must Sample Sizes Be Equal? 400
Using Software for Calculations 400
Analogy to Confidence Intervals 400
Separate Confidence Intervals 401
Testing for Nonzero Difference (Optional) 403
Test Statistic 403
Using the p-Value 404

10.3 Comparing Two Means: Independent Samples 406

Format of Hypotheses 406
Test Statistic 406
Case 1: Known Variances 407
Case 2: Unknown Variances, Assumed Equal 407
Case 3: Unknown Variances, Assumed Unequal 407
Which Assumption is Best? 411
Must Sample Sizes Be Equal? 411
Large Samples 411
Caution: Three Issues 411

10.4 Comparing Two Means: Paired Samples 414

Paired Data 414
Paired t Test 414
Excel's Paired Difference Test 416
Analogy to Confidence Interval 416
Why Not Treat Paired Data As Independent Samples? 417

10.5 Comparing Two Variances 420

Format of Hypotheses 420
The F Test 420
Critical Values 420
Illustration: Collision Damage 421
Comparison of Means 422
Comparison of Variances: Two-Tailed Test 422
Comparison of Variances: One-Tailed Test 424
Excel's F Test 425
Assumptions of the F Test 425
Significance versus Importance 426
 Chapter Summary 427

CHAPTER ELEVEN

Analysis of Variance 438

11.1 Overview of ANOVA 439

The Goal: Explaining Variation 439
Illustration: Manufacturing Defect Rates 440
Illustration: Hospital Length of Stay 441
Illustration: Automobile Painting 441
ANOVA Calculations 441
ANOVA Assumptions 442

11.2 One-Factor ANOVA (Completely Randomized Model) 442

Data Format 442
Hypotheses to Be Tested 443
One-Factor ANOVA as a Linear Model 443
Group Means 443
Partitioned Sum of Squares 443

Test Statistic 445
Decision Rule 445
Using MINITAB 448

11.3 Multiple Comparisons 450

Tukey's Test 450
Using MegaStat 451

11.4 Tests for Homogeneity of Variances (Optional) 452

ANOVA Assumptions 452
Hartley's F_{\max} Test 452
Levene's Test 454

11.5 Two-Factor ANOVA without Replication (Randomized Block Model) 456

Data Format 456
Two-Factor ANOVA Model 457
Hypotheses to Be Tested 457
Randomized Block Model 457
Format of Calculation of Nonreplicated Two-Factor ANOVA 458
Using MegaStat 460
Multiple Comparisons 461
Limitations of Two-Factor ANOVA without Replication 461

11.6 Two-Factor ANOVA with Replication (Full Factorial Model) 464

What Does Replication Accomplish? 464
Format of Hypotheses 464
Format of Data 465
Sources of Variation 465
Using MegaStat 468
Interaction Effect 468
Tukey Tests of Pairs of Means 470
Significance versus Importance 470

11.7 General Linear Model (Optional) 473

Higher-Order ANOVA Models 473
What Is GLM? 474

11.8 Experimental Design: An Overview (Optional) 476

What Is Experimental Design? 476
 2^k Models 476
Fractional Factorial Designs 476
Nested or Hierarchical Design 477
Random Effects Models 477
 Chapter Summary 477

CHAPTER TWELVE

Bivariate Regression 488

12.1 Visual Displays and Correlation Analysis 489

Visual Displays 489
Correlation Coefficient 490
Tests for Significance 490
Quick Rule for Significance 494
Role of Sample Size 494
Using Excel 494
Regression: The Next Step? 496
Autocorrelation 499

- 12.2 Bivariate Regression** 500
What Is Bivariate Regression? 500
Model Form 500
Interpreting a Fitted Regression 501
Prediction Using Regression 501
- 12.3 Regression Terminology** 502
Models and Parameters 502
Estimating a Regression Line by Eye 502
Fitting a Regression on a Scatter Plot in Excel 502
Illustration: Piper Cheyenne Fuel Consumption 503
- 12.4 Ordinary Least Squares Formulas** 505
Slope and Intercept 505
Illustration: Exam Scores and Study Time 506
Assessing Fit 507
Coefficient of Determination 508
- 12.5 Tests for Significance** 511
Standard Error of Regression 511
Confidence Intervals for Slope and Intercept 511
Hypothesis Tests 512
Test for Zero Slope: Exam Scores 512
Using Excel: Exam Scores 513
Using MegaStat: Exam Scores 513
Using MINITAB: Exam Scores 514
Using MegaStat: U.S. Income and Taxes 516
MegaStat's Confidence Intervals: U.S. Income and Taxes 516
Test for Zero Slope: Tax Data 516
- 12.6 Analysis of Variance: Overall Fit** 517
Decomposition of Variance 517
F Statistic for Overall Fit 518
- 12.7 Confidence and Prediction Intervals for Y** 522
How to Construct an Interval Estimate for Y 522
Two Illustrations: Exam Scores and Taxes 522
Quick Rules for Confidence and Prediction Intervals 524
- 12.8 Violations of Assumptions** 524
Three Important Assumptions 524
Non-Normal Errors 524
Histogram of Residuals 524
Normal Probability Plot 525
What to Do About Non-Normality? 525
Heteroscedastic Errors (Nonconstant Variance) 526
Tests for Heteroscedasticity 526
What to Do about Heteroscedasticity? 527
Autocorrelated Errors 527
Runs Test for Autocorrelation 528
Durbin-Watson Test 528
What to Do about Autocorrelation? 529
- 12.9 Unusual Observations** 531
Standardized Residuals: Excel 531
Studentized Residuals: MINITAB 531
Studentized Residuals: MegaStat 532
Leverage and Influence 532
Studentized Deleted Residuals 534
- 12.10 Other Regression Problems (Optional)** 536
Outliers 536
Model Misspecification 536
Ill-Conditioned Data 536
Spurious Correlation 537
Model Form and Variable Transforms 538
Regression by Splines 540
Chapter Summary 541
- CHAPTER THIRTEEN**
Multiple Regression 558
- 13.1 Multiple Regression** 559
Bivariate or Multivariate? 559
Regression Terminology 560
Data Format 560
Illustration: Home Prices 560
Logic of Variable Selection 561
Fitted Regression 561
Two-Predictor Model 562
One-Predictor Model 562
Common Misconceptions about Fit 562
Regression Modeling 563
- 13.2 Assessing Overall Fit** 564
F Test for Significance 564
Coefficient of Determination (R^2) 565
Adjusted R^2 565
How Many Predictors? 565
- 13.3 Predictor Significance** 566
Hypothesis Tests 566
Test Statistic 567
- 13.4 Confidence Intervals for Y** 569
Standard Error 569
Approximate Confidence and Prediction Intervals for Y 569
Quick 95 Percent Prediction Interval for Y 570
- 13.5 Binary Predictors** 572
What Is a Binary Predictor? 572
Effects of a Binary Predictor 572
Testing a Binary for Significance 573
More Than One Binary 574
What If I Forget to Exclude One Binary? 575
Regional Binaries 577
- 13.6 Tests for Nonlinearity and Interaction** 579
Tests for Nonlinearity 579
Tests for Interaction 580
- 13.7 Multicollinearity** 581
What Is Multicollinearity? 581
Variance Inflation 581
Correlation Matrix 582
Predictor Matrix Plots 582
Variance Inflation Factor (VIF) 583
Rules of Thumb 583
Are Coefficients Stable? 584
- 13.8 Violations of Assumptions** 586
Non-Normal Errors 586

Nonconstant Variance (Heteroscedasticity) 586
Autocorrelation (Optional) 588
Unusual Observations 588

13.9 Other Regression Topics 590

Outliers: Causes and Cures 590
Missing Predictors 590
Ill-Conditioned Data 590
Significance in Large Samples 591
Model Specification Errors 591
Missing Data 591
Binary Dependent Variable 591
Stepwise and Best Subsets Regression 591

Chapter Summary 592

CHAPTER FOURTEEN Time-Series Analysis 604

14.1 Time-Series Components 605

Time-Series Data 605
Stocks and Flows 606
Periodicity 607
Additive versus Multiplicative Models 607
A Graphical View 607
Trend 607
Cycle 609
Seasonal 609
Irregular 609

14.2 Trend Forecasting 610

Three Trend Models 610
Linear Trend Model 610
Illustration: Linear Trend 611
Linear Trend Calculations 611
Forecasting a Linear Trend 612
Linear Trend: Calculating R^2 612
Exponential Trend Model 612
When to Use the Exponential Model 613
Illustration: Exponential Trend 613
Exponential Trend Calculations 614
Forecasting an Exponential Trend 615
Exponential Trend: Calculating R^2 615
Quadratic Trend Model 615
Illustration: Quadratic Trend 616
Using Excel for Trend Fitting 617
Trend-Fitting Criteria 617

14.3 Assessing Fit 623

Five Measures of Fit 623

14.4 Moving Averages 625

Trendless or Erratic Data 625
Trailing Moving Average (TMA) 625
Centered Moving Average (CMA) 626
Using Excel for a TMA 626

14.5 Exponential Smoothing 627

Forecast Updating 627
Smoothing Constant (α) 628
Choosing the Value of α 628
Initializing the Process 628
Using MINITAB 630

Using Excel 631

Smoothing with Trend and Seasonality 631

14.6 Seasonality 633

When and How to Deseasonalize 633
Illustration of Calculations 633
Using MINITAB to Deseasonalize 635
Seasonal Forecasts Using Binary Predictors 636

14.7 Forecasting: Final Thoughts 639

Role of Forecasting 639
Behavioral Aspects of Forecasting 639
Forecasts Are Always Wrong 639

Chapter Summary 640

CHAPTER FIFTEEN Chi-Square Tests 656

15.1 Chi-Square Test for Independence 657

Contingency Tables 657
Chi-Square Test 658
Chi-Square Distribution 658
Expected Frequencies 659
Illustration of the Chi-Square Calculations 660
Test of Two Proportions 662
Small Expected Frequencies 663
Cross-Tabulating Raw Data 663
3-Way Tables and Higher 664

15.2 Chi-Square Tests for Goodness-of-Fit 667

Purpose of the Test 667
Hypotheses for GOF 667
Test Statistic and Degrees of Freedom for GOF 667
Data-Generating Situations 668
Mixtures: A Problem 668
Eyeball Tests 668
Small Expected Frequencies 668

15.3 Uniform Goodness-of-Fit Test 669

Multinomial Distribution 669
Uniform Distribution 669
Uniform GOF Test: Grouped Data 669
Uniform GOF Test: Raw Data 670

15.4 Poisson Goodness-of-Fit Test 673

Poisson Data-Generating Situations 673
Poisson Goodness-of-Fit Test 674
Poisson GOF Test: Tabulated Data 674
Poisson GOF Test: Raw Data 676

15.5 Normal Chi-Square Goodness-of-Fit Test 679

Normal Data-Generating Situations 679
Method 1: Standardizing the Data 679
Method 2: Equal Bin Widths 679
Method 3: Equal Expected Frequencies 680
Application: Quality Management 680

15.6 ECDF Tests (Optional) 684

Kolmogorov-Smirnov and Lilliefors Tests 684
Illustrations: Lottery Numbers and Kiss Weights 684
Anderson-Darling Test 685

Chapter Summary 686

CHAPTER SIXTEEN**Nonparametric Tests 698**

- 16.1** Why Use Nonparametric Tests? 699
- 16.2** One-Sample Runs Test 700
 - Application: Quality Inspection* 700
 - Small Samples* 702
- 16.3** Wilcoxon Signed-Rank Test 702
 - Application: Median versus Benchmark* 703
 - Application: Paired Data* 704
- 16.4** Mann-Whitney Test 706
 - Application: Restaurant Quality* 706
- 16.5** Kruskal-Wallis Test for Independent Samples 709
 - Application: Employee Absenteeism* 709
- 16.6** Friedman Test for Related Samples 714
 - Test Statistic* 714
 - Application: Braking Effectiveness* 714
- 16.7** Spearman Rank Correlation Test 716
 - Application: Calories and Fat* 717
 - Correlation versus Causation* 718
- Chapter Summary 720

CHAPTER SEVENTEEN**Quality Management 730**

- 17.1** Quality and Variation 731
 - What Is Quality?* 731
 - Productivity and Quality* 732
 - Processes and Quality Metrics* 732
 - Variance Reduction* 732
 - Common Cause versus Special Cause* 733
 - Role of Management* 733
 - Role of Statisticians* 733
- 17.2** Customer Orientation 733
 - Who Is a Customer?* 733
 - Measuring Quality* 734
- 17.3** Behavioral Aspects of Quality 734
 - Blame versus Solutions* 734
 - Employee Involvement* 735
- 17.4** Pioneers in Quality Management 735
 - Brief History of Quality Control* 735
 - W. Edwards Deming* 736
 - Other Influential Thinkers* 737
- 17.5** Quality Improvement 737
 - Total Quality Management (TQM)* 737
 - Business Process Redesign (BPR)* 738
 - Statistical Quality Control (SQC)* 738
 - Statistical Process Control (SPC)* 738
 - Continuous Quality Improvement (CQI)* 739
- 17.6** Control Charts: Overview 740
 - What Is a Control Chart?* 740
 - Two Data Types* 740
 - Three Common Control Charts* 740
- 17.7** Control Charts for a Mean 741
 - \bar{x} Charts: Bottle-Filling Example* 741
 - Control Limits: Known μ and σ* 741

Empirical Control Limits 743
Control Chart Factors 743
Detecting Abnormal Patterns 744
Histograms 746

- 17.8** Control Charts for a Range 749
 - Control Limits for the Range* 749
- 17.9** Patterns in Control Charts 750
 - The Overadjustment Problem* 750
 - Abnormal Patterns* 750
 - Symptoms and Assignable Causes* 751
- 17.10** Process Capability 752
 - C_p Index* 752
 - C_{pk} Index* 753
 - Bottle Filling Revisited* 755
- 17.11** Other Control Charts 755
 - Attribute Data: p Charts* 755
 - Application: Emergency Patients* 757
 - Other Standard Control Charts (s , c , np , I , MR)* 757
 - Ad Hoc Charts* 759
- 17.12** Additional Quality Topics (Optional) 760
 - Acceptance Sampling* 760
 - Supply-Chain Management* 760
 - Quality and Design* 761
 - Taguchi's Robust Design* 761
 - Six Sigma and Lean Six Sigma* 761
 - ISO 9000* 762
 - Malcolm Baldrige Award* 762
 - Advanced MINITAB Features* 762
 - Future of Statistical Process Control* 762
- Chapter Summary 763

CHAPTER EIGHTEEN**Simulation (On Student CD-ROM)****APPENDIXES**

- A** Exact Binomial Probabilities 774
- B** Exact Poisson Probabilities 776
- C-1** Standard Normal Areas 779
- C-2** Cumulative Standard Normal Distribution 780
- D** Student's t Critical Values 782
- E** Chi-Square Critical Values 783
- F** Critical Values of F 784
- G** Solutions to Odd-Numbered Exercises 792

PHOTO CREDITS 816**INDEX 817**