The PRACTICE *of* STATISTICS *for* BUSINESS *and* ECONOMICS

FOURTH EDITION

David S. Moore
Purdue University

George P. McCabe
Purdue University

Layth C. Alwan
University of Wisconsin–Milwaukee

Bruce A. Craig
Purdue University

W. H. Freeman and Company
A Macmillan Education Imprint

BRIEF CONTENTS

CHAPTER 1	Examining Distributions	1
CHAPTER 2	Examining Relationships	63
CHAPTER 3	Producing Data	123
CHAPTER 4	Probability: The Study of Randomness	173
CHAPTER 5	Distributions for Counts and Proportions	243
CHAPTER 6	Introduction to Inference	287
CHAPTER 7	Inference for Means	357
CHAPTER 8	Inference for Proportions	417
CHAPTER 9	Inference for Categorical Data	455
CHAPTER 10	Inference for Regression	483
CHAPTER 11	Multiple Regression	531
CHAPTER 12	Statistics for Quality: Control and Capability	591
CHAPTER 13	Time Series Forecasting	643
CHAPTER 14	One-Way Analysis of Variance	711

The Core book includes Chapters 1–14. Chapters 15–17 are individual optional Companion Chapters and can be found at www.macmillanhighered.com/psbe4e.

CHAPTER 15	Two-Way Analysis of Variance	15-1
CHAPTER 16	Nonparametric Tests	16-1
CHAPTER 17	Logistic Regression	17-1

CONTENTS

To Instructors: About This Book xv
Media and Supplements xxi
To Students: What Is Statistics? xxiv
Index of Cases xxviii
Index of Data Tables xxix
Beyond the Basics Index xxx
About the Authors xxxi

CHAPTER 1 Examining Distributions 1

Introduction 1

1.1 Data 2
Section 1.1 Summary 6
Section 1.1 Exercises 6

1.2 Displaying Distributions with Graphs 7
Categorical variables: Bar graphs and pie charts 7
Quantitative variables: Histograms 12
Case 1.1 Treasury Bills 12
Quantitative variables: Stemplots 15
Interpreting histograms and stemplots 17
Time plots 19
Section 1.2 Summary 20
Section 1.2 Exercises 21

1.3 Describing Distributions with Numbers 23
Case 1.2 Time to Start a Business 23
Measuring center: The mean 24
Measuring center: The median 25
Comparing the mean and the median 26
Measuring spread: The quartiles 27
The five-number summary and boxplots 29
Measuring spread: The standard deviation 31
Choosing measures of center and spread 32
BEYOND THE BASICS: *Risk and Return* 33
Section 1.3 Summary 34
Section 1.3 Exercises 35

1.4 Density Curves and the Normal Distributions 38
Density curves 38
The median and mean of a density curve 39
Normal distributions 42
The 68–95–99.7 rule 43
The standard Normal distribution 45
Normal distribution calculations 46
Using the standard Normal table 48
Inverse Normal calculations 49
Assessing the Normality of data 51
BEYOND THE BASICS: *Density Estimation* 54
Section 1.4 Summary 56
Section 1.4 Exercises 57

Chapter 1 Review Exercises 59

CHAPTER 2 Examining Relationships 63

Introduction 63

2.1 Scatterplots 65
Case 2.1 Education Expenditures and Population: Benchmarking 65
Interpreting scatterplots 67
The log transformation 68
Adding categorical variables to scatterplots 70
Section 2.1 Summary 71
Section 2.1 Exercises 72

2.2 Correlation 74
The correlation *r* 75
Facts about correlation 76
Section 2.2 Summary 78
Section 2.2 Exercises 78

2.3 Least-Squares Regression 80
The least-squares regression line 81
Facts about least-squares regression 86
Interpretation of r^2 87
Residuals 88
The distribution of the residuals 92
Influential observations 92
Section 2.3 Summary *94*
Section 2.3 Exercises *95*

2.4 Cautions about Correlation and Regression 98
Extrapolation 98
Correlations based on averaged data 99
Lurking variables 100
Association is not causation 101
BEYOND THE BASICS: *Data Mining* ***102***
Section 2.4 Summary *103*
Section 2.4 Exercises *103*

2.5 Relations in Categorical Data 104
Case 2.2 Does the Right Music Sell the Product? 104
Marginal distributions 105
Conditional distributions 107
Mosaic plots and software output 109
Simpson's paradox 110
Section 2.5 Summary *112*
Section 2.5 Exercises *113*

Chapter 2 Review Exercises 116

CHAPTER 3 Producing Data 123

Introduction 123

3.1 Sources of Data 124
Anecdotal data 124
Available data 124
Sample surveys and experiments 126
Section 3.1 Summary *127*
Section 3.1 Exercises *128*

3.2 Designing Samples 129
Simple random samples 132
Stratified samples 134
Multistage samples 135
Cautions about sample surveys 136
BEYOND THE BASICS: *Capture-Recapture Sampling* ***139***
Section 3.2 Summary *140*
Section 3.2 Exercises *140*

3.3 Designing Experiments 142
Comparative experiments 145
Randomized comparative experiments 146
Completely randomized designs 147
How to randomize 147
The logic of randomized comparative experiments 150
Cautions about experimentation 153
Matched pairs designs 154
Block designs 155
Section 3.3 Summary *157*
Section 3.3 Exercises *157*

3.4 Data Ethics 160
Institutional review boards 160
Informed consent 161
Confidentiality 162
Clinical trials 163
Behavioral and social science experiments 165
Section 3.4 Summary *167*
Section 3.4 Exercises *167*

Chapter 3 Review Exercises 169

CHAPTER 4 Probability: The Study of Randomness 173

Introduction 173

4.1 Randomness 174
The language of probability 175
Thinking about randomness and probability 176
Section 4.1 Summary *176*
Section 4.1 Exercises *177*

4.2 Probability Models 179
Sample spaces 179
Probability rules 182
Assigning probabilities: Finite number of outcomes 184
Case 4.1 Uncovering Fraud by Digital Analysis 184
Assigning probabilities: Equally likely outcomes 186

Independence and the multiplication rule 187
Applying the probability rules 189
Section 4.2 Summary *191*
Section 4.2 Exercises *191*

4.3 General Probability Rules 194
General addition rules 194
Conditional probability 197
General multiplication rules 200
Tree diagrams 201
Bayes's rule 203
Independence again 205
Section 4.3 Summary *205*
Section 4.3 Exercises *206*

4.4 Random Variables 209
Discrete random variables 210
Case 4.2 Tracking Perishable Demand 210
Continuous random variables 213
Normal distributions as probability distributions 215
Section 4.4 Summary *216*
Section 4.4 Exercises *217*

4.5 Means and Variances of Random Variables 219
The mean of a random variable 219
Mean and the law of large numbers 222
Thinking about the law of large numbers 223
Rules for means 224
Case 4.3 Portfolio Analysis 225
The variance of a random variable 229
Rules for variances and standard deviations 230
Section 4.5 Summary *235*
Section 4.5 Exercises *236*
Chapter 4 Review Exercises 239

CHAPTER 5 Distributions for Counts and Proportions 243

Introduction 243

5.1 The Binomial Distributions 244
The binomial distributions for sample counts 245
The binomial distributions for statistical sampling 247
Case 5.1 Inspecting a Supplier's Products 247
Finding binomial probabilities 247
Binomial formula 250
Binomial mean and standard deviation 253
Sample proportions 255
Normal approximation for counts and proportions 256
The continuity correction 260
Assessing binomial assumption with data 261
Section 5.1 Summary *263*
Section 5.1 Exercises *264*

5.2 The Poisson Distributions 267
The Poisson setting 267
The Poisson model 269
Approximations to the Poisson 270
Assessing Poisson assumption with data 271
Section 5.2 Summary *273*
Section 5.2 Exercises *273*

5.3 Toward Statistical Inference 275
Sampling distributions 276
Bias and variability 279
Why randomize? 280
Section 5.3 Summary *281*
Section 5.3 Exercises *281*
Chapter 5 Review Exercises 284

CHAPTER 6 Introduction to Inference 287

Introduction 287

Overview of inference 288

6.1 The Sampling Distribution of a Sample Mean 288
The mean and standard deviation of $\bar{x}$ 292
The central limit theorem 294
Section 6.1 Summary *299*
Section 6.1 Exercises *300*

6.2 Estimating with Confidence 302
Statistical confidence 302
Confidence intervals 304
Confidence interval for a population mean 306
How confidence intervals behave 309
Some cautions 311
Section 6.2 Summary *313*
Section 6.2 Exercises *314*

6.3 Tests of Significance 316
The reasoning of significance tests 316
Case 6.1 Fill the Bottles 317
Stating hypotheses 319
Test statistics 321
P-values 322
Statistical significance 324
Tests of one population mean 326
Two-sided significance tests and confidence intervals 329
P-values versus reject-or-not reporting 331
Section 6.3 Summary *332*
Section 6.3 Exercises *333*

6.4 Using Significance Tests 336
Choosing a level of significance 337
What statistical significance does not mean 337
Statistical inference is not valid for all sets of data 339
Beware of searching for significance 339
Section 6.4 Summary *341*
Section 6.4 Exercises *341*

6.5 Power and Inference as a Decision 343
Power 343
Increasing the power 345
Inference as decision 346
Two types of error 346
Error probabilities 347
The common practice of testing hypotheses 349
Section 6.5 Summary *350*
Section 6.5 Exercises *350*

Chapter 6 Review Exercises 351

CHAPTER 7 Inference for Means 357

Introduction 357

7.1 Inference for the Mean of a Population 358
t distributions 358
The one-sample *t* confidence interval 360
Case 7.1 Time Spent Using a Smartphone 361
The one-sample *t* test 362
Using software 365
Matched pairs *t* procedures 368
Robustness of the one-sample *t* procedures 371
BEYOND THE BASICS: The Bootstrap 372
Section 7.1 Summary *373*
Section 7.1 Exercises *374*

7.2 Comparing Two Means 378
The two-sample *t* statistic 379
The two-sample *t* confidence interval 381
The two-sample *t* significance test 383
Robustness of the two-sample procedures 384
Inference for small samples 385
The pooled two-sample *t* procedures 386
Case 7.2 Active versus Failed Retail Companies 389
Section 7.2 Summary *392*
Section 7.2 Exercises *393*

7.3 Additional Topics on Inference 398
Choosing the sample size 398
Inference for non-Normal populations 406
Section 7.3 Summary *408*
Section 7.3 Exercises *409*

Chapter 7 Review Exercises 411

CHAPTER 8 Inference for Proportions 417

Introduction 417

8.1 Inference for a Single Proportion 418

Case 8.1 Robotics and Jobs 418
Large-sample confidence interval for a single proportion 420
Plus four confidence interval for a single proportion 421
Significance test for a single proportion 423
Choosing a sample size for a confidence interval 426
Case 8.2 Marketing Christmas Trees 428
Choosing a sample size for a significance test 429
Section 8.1 Summary *431*
Section 8.1 Exercises *432*

8.2 Comparing Two Proportions 436
Large-sample confidence intervals for a difference in proportions 437
Case 8.3 Social Media in the Supply Chain 438
Plus four confidence intervals for a difference in proportions 440
Significance tests 440
Choosing a sample size for two sample proportions 444
BEYOND THE BASICS: *Relative Risk* *447*
Section 8.2 Summary *448*
Section 8.2 Exercises *449*
Chapter 8 Review Exercises 451

CHAPTER 9 Inference for Categorical Data 455

9.1 Inference for Two-Way Tables 455
Two-way tables 456
Case 9.1 Are Flexible Companies More Competitive? 457
Describing relations in two-way tables 458
The hypothesis: No association 462
Expected cell counts 462
The chi-square test 463
The chi-square test and the z test 465
Models for two-way tables 466
BEYOND THE BASICS: *Meta-Analysis* *468*
Section 9.1 Summary *470*

9.2 Goodness of Fit 470
Section 9.2 Summary *475*
Chapter 9 Review Exercises 475

CHAPTER 10 Inference for Regression 483

Introduction 483

10.1 Inference about the Regression Model 484
Statistical model for simple linear regression 484
From data analysis to inference 485
Case 10.1 The Relationship between Income and Education for Entrepreneurs 485
Estimating the regression parameters 490
Conditions for regression inference 494
Confidence intervals and significance tests 495
The word "regression" 500
Inference about correlation 500
Section 10.1 Summary *502*
Section 10.1 Exercises *503*

10.2 Using the Regression Line 510
BEYOND THE BASICS: *Nonlinear Regression* *515*
Section 10.2 Summary *515*
Section 10.2 Exercises *516*

10.3 Some Details of Regression Inference 517
Standard errors 518
Analysis of variance for regression 520
Section 10.3 Summary *524*
Section 10.3 Exercises *524*
Chapter 10 Review Exercises 526

CHAPTER 11 Multiple Regression 531

Introduction 531

11.1 Data Analysis for Multiple Regression 534
Case 11.1 Assets, Sales, and Profits 534
Data for multiple regression 534
Preliminary data analysis for multiple regression 535
Estimating the multiple regression coefficients 538
Regression residuals 541
The regression standard error 544
Section 11.1 Summary 545
Section 11.1 Exercises 546

11.2 Inference for Multiple Regression 548
Multiple linear regression model 549
Case 11.2 Predicting Movie Revenue 550
Estimating the parameters of the model 550
Inference about the regression coefficients 551
Inference about prediction 554
ANOVA table for multiple regression 555
Squared multiple correlation R^2 558
Inference for a collection of regression coefficients 559
Section 11.2 Summary 561
Section 11.2 Exercises 562

11.3 Multiple Regression Model Building 566
Case 11.3 Prices of Homes 566
Models for curved relationships 569
Models with categorical explanatory variables 571
More elaborate models 575
Variable selection methods 577
BEYOND THE BASICS: *Multiple Logistic Regression 580*
Section 11.3 Summary 582
Section 11.3 Exercises 582

Chapter 11 Review Exercises 584

CHAPTER 12 Statistics for Quality: Control and Capability 591

Introduction 591
Quality overview 592
Systematic approach to process improvement 593
Process improvement toolkit 594

12.1 Statistical Process Control 597
Section 12.1 Summary 599
Section 12.1 Exercises 600

12.2 Variable Control Charts 600
$\bar{x}$ and R charts 601
Case 12.1 Turnaround Time for Lab Results 604
Case 12.2 O-Ring Diameters 609
$\bar{x}$ and s charts 612
Charts for individual observations 614
Don't confuse control with capability! 619
Section 12.2 Summary 626
Section 12.2 Exercises 626

12.3 Attribute Control Charts 630
Control charts for sample proportions 630
Case 12.3 Reducing Absenteeism 631
Control charts for counts per unit of measure 636
Section 12.3 Summary 638
Section 12.3 Exercises 638

Chapter 12 Review Exercises 639

CHAPTER 13 Time Series Forecasting 643

Introduction 643
Overview of Time Series Forecasting 643

13.1 Assessing Time Series Behavior 644
Case 13.1 Amazon Sales 647
Runs test 648
Autocorrelation function 651

Forecasts 656
Section 13.1 Summary *657*
Section 13.1 Exercises *657*

13.2 Random Walks 657
Price changes versus returns 661
Section 13.2 Summary *663*
Section 13.2 Exercises *663*

13.3 Modeling Trend and Seasonality Using Regression 664
Identifying trends 665
Seasonal patterns 671
Using indicator variables 672
Residual checking 677
Section 13.3 Summary *679*
Section 13.3 Exercises *679*

13.4 Lag Regression Models 681
Autoregressive-based models 681
Section 13.4 Summary *690*
Section 13.4 Exercises *691*

13.5 Moving-Average and Smoothing Models 691
Moving-average models 691
Moving average and seasonal ratios 694
Exponential smoothing models 699
Section 13.5 Summary *706*
Section 13.5 Exercises *706*

Chapter 13 Review Exercises 708

CHAPTER 14 One-Way Analysis of Variance 711

Introduction 711

14.1 One-Way Analysis of Variance 712
The ANOVA setting 712
Comparing means 713
The two-sample *t* statistic 714
An overview of ANOVA 715
Case 14.1 Tip of the Hat and Wag of the Finger? 715
The ANOVA model 718
Estimates of population parameters 720
Testing hypotheses in one-way ANOVA 722
The ANOVA table 724
The *F* test 726
Using software 729
BEYOND THE BASICS: *Testing the Equality of Spread* *731*
Section 14.1 Summary *733*

14.2 Comparing Group Means 733
Contrasts 733
Case 14.2 Evaluation of a New Educational Product 733
Multiple comparisons 740
Simultaneous confidence intervals 744
Section 14.2 Summary *744*

14.3 The Power of the ANOVA Test 745
Section 14.3 Summary *749*

Chapter 14 Review Exercises 749

Notes and Data Sources *N-1*
Tables *T-1*
Answers to Odd-Numbered Exercises *S-1*
Index *I-1*

The following optional Companion Chapters can be found online at **www.macmillanhighered.com/psbe4e.**

CHAPTER 15 Two-Way Analysis of Variance 15-1

Introduction 15-1

15.1 The Two-Way ANOVA Model 15-2
Advantages of two-way ANOVA 15-2
The two-way ANOVA model 15-5
Main effects and interactions 15-7
Section 15.1 Summary *15-13*

15.2 Inference for Two-Way ANOVA 15-13
The ANOVA table for two-way ANOVA 15-13
Carrying out a two-way ANOVA 15-15
Case 15.1 Discounts and Expected Prices 15-15

Case 15.2 Expected Prices, Continued 15-16
Section 15.2 Summary *15-21*
Chapter 15 Review Exercises 15-21

Notes and Data Sources *15-28*
Answers to Odd-Numbered Exercises *15-29*

CHAPTER 16 Nonparametric Tests 16-1

Introduction 16-1

16.1 The Wilcoxon Rank Sum Test 16-3
Case 16.1 Price Discrimination? 16-3
The rank transformation 16-4
The Wilcoxon rank sum test 16-5
The Normal approximation 16-7
What hypotheses do the Wilcoxon test? 16-8
Ties 16-9
Case 16.2 Consumer Perceptions of Food Safety 16-10
Rank versus *t* tests 16-12
Section 16.1 Summary *16-12*
Section 16.1 Exercises *16-12*

16.2 The Wilcoxon Signed Rank Test 16-15
The Normal approximation 16-18
Ties 16-19
Section 16.2 Summary *16-21*
Section 16.2 Exercises *16-21*

16.3 The Kruskal-Wallis Test 16-24
Hypotheses and assumptions 16-25
The Kruskal-Wallis test 16-25
Section 16.3 Summary *16-27*
Section 16.3 Exercises *16-27*
Chapter 16 Review Exercises 16-29

Notes and Data Sources *16-31*
Answers to Odd-Numbered Exercises *16-32*

CHAPTER 17 Logistic Regression 17-1

Introduction 17-1

17.1 The Logistic Regression Model 17-2
Case 17.1 Clothing Color and Tipping 17-2
Binomial distributions and odds 17-2
Model for logistic regression 17-4
Fitting and interpreting the logistic regression model 17-5
Section 17.1 Summary *17-8*

17.2 Inference for Logistic Regression 17-9
Examples of logistic regression analyses 17-11
Section 17.2 Summary *17-15*

17.3 Multiple Logistic Regression 17-16
Section 17.3 Summary *17-17*
Chapter 17 Review Exercises 17-17

Notes and Data Sources *17-22*
Answers to Odd-Numbered Exercises *17-23*