

Table of Contents

PREFACE	v
LIST OF EXAMPLES	xxi
1 An Introduction to Econometrics	1
1.1 Why Study Econometrics?	1
1.2 What Is Econometrics About?	2
1.2.1 Some Examples	3
1.3 The Econometric Model	4
1.3.1 Causality and Prediction	5
1.4 How Are Data Generated?	5
1.4.1 Experimental Data	6
1.4.2 Quasi-Experimental Data	6
1.4.3 Nonexperimental Data	7
1.5 Economic Data Types	7
1.5.1 Time-Series Data	7
1.5.2 Cross-Section Data	8
1.5.3 Panel or Longitudinal Data	9
1.6 The Research Process	9
1.7 Writing an Empirical Research Paper	11
1.7.1 Writing a Research Proposal	11
1.7.2 A Format for Writing a Research Report	11
1.8 Sources of Economic Data	13
1.8.1 Links to Economic Data on the Internet	13
1.8.2 Interpreting Economic Data	14
1.8.3 Obtaining the Data	14
Probability Primer	15
P.1 Random Variables	16
P.2 Probability Distributions	17
P.3 Joint, Marginal, and Conditional Probabilities	20
P.3.1 Marginal Distributions	20
P.3.2 Conditional Probability	21
P.3.3 Statistical Independence	21
P.4 A Digression: Summation Notation	22
P.5 Properties of Probability Distributions	23
P.5.1 Expected Value of a Random Variable	24
P.5.2 Conditional Expectation	25
P.5.3 Rules for Expected Values	25
P.5.4 Variance of a Random Variable	26
P.5.5 Expected Values of Several Random Variables	27
P.5.6 Covariance Between Two Random Variables	27
P.6 Conditioning	29
P.6.1 Conditional Expectation	30
P.6.2 Conditional Variance	31
P.6.3 Iterated Expectations	32
P.6.4 Variance Decomposition	33
P.6.5 Covariance Decomposition	34
P.7 The Normal Distribution	34
P.7.1 The Bivariate Normal Distribution	37
P.8 Exercises	39
2 The Simple Linear Regression Model	46
2.1 An Economic Model	47
2.2 An Econometric Model	49
2.2.1 Data Generating Process	51
2.2.2 The Random Error and Strict Exogeneity	52
2.2.3 The Regression Function	53
2.2.4 Random Error Variation	54
2.2.5 Variation in x	56
2.2.6 Error Normality	56
2.2.7 Generalizing the Exogeneity Assumption	56
2.2.8 Error Correlation	57
2.2.9 Summarizing the Assumptions	58
2.3 Estimating the Regression Parameters	59
2.3.1 The Least Squares Principle	61
2.3.2 Other Economic Models	65
2.4 Assessing the Least Squares Estimators	66
2.4.1 The Estimator b_2	67
2.4.2 The Expected Values of b_1 and b_2	68
2.4.3 Sampling Variation	69
2.4.4 The Variances and Covariance of b_1 and b_2	69
2.5 The Gauss–Markov Theorem	72
2.6 The Probability Distributions of the Least Squares Estimators	73
2.7 Estimating the Variance of the Error Term	74
2.7.1 Estimating the Variances and Covariance of the Least Squares Estimators	74
2.7.2 Interpreting the Standard Errors	76

2.8	Estimating Nonlinear Relationships	77	3.2	Hypothesis Tests	118
2.8.1	Quadratic Functions	77	3.2.1	The Null Hypothesis	118
2.8.2	Using a Quadratic Model	77	3.2.2	The Alternative Hypothesis	118
2.8.3	A Log-Linear Function	79	3.2.3	The Test Statistic	119
2.8.4	Using a Log-Linear Model	80	3.2.4	The Rejection Region	119
2.8.5	Choosing a Functional Form	82	3.2.5	A Conclusion	120
2.9	Regression with Indicator Variables	82	3.3	Rejection Regions for Specific Alternatives	120
2.10	The Independent Variable	84	3.3.1	One-Tail Tests with Alternative “Greater Than” ($>$)	120
2.10.1	Random and Independent x	84	3.3.2	One-Tail Tests with Alternative “Less Than” ($<$)	121
2.10.2	Random and Strictly Exogenous x	86	3.3.3	Two-Tail Tests with Alternative “Not Equal To” (\neq)	122
2.10.3	Random Sampling	87	3.4	Examples of Hypothesis Tests	123
2.11	Exercises	89	3.5	The p-Value	126
2.11.1	Problems	89	3.6	Linear Combinations of Parameters	129
2.11.2	Computer Exercises	93	3.6.1	Testing a Linear Combination of Parameters	131
Appendix 2A	Derivation of the Least Squares Estimates	98	3.7	Exercises	133
Appendix 2B	Deviation from the Mean Form of b_2	99	3.7.1	Problems	133
Appendix 2C	b_2 Is a Linear Estimator	100	3.7.2	Computer Exercises	139
Appendix 2D	Derivation of Theoretical Expression for b_2	100	Appendix 3A	Derivation of the t-Distribution	144
Appendix 2E	Deriving the Conditional Variance of b_2	100	Appendix 3B	Distribution of the t-Statistic under H_1	145
Appendix 2F	Proof of the Gauss–Markov Theorem	102	Appendix 3C	Monte Carlo Simulation	147
Appendix 2G	Proofs of Results Introduced in Section 2.10	103	3C.1	Sampling Properties of Interval Estimators	148
2G.1	The Implications of Strict Exogeneity	103	3C.2	Sampling Properties of Hypothesis Tests	149
2G.2	The Random and Independent x Case	103	3C.3	Choosing the Number of Monte Carlo Samples	149
2G.3	The Random and Strictly Exogenous x Case	105	3C.4	Random- x Monte Carlo Results	150
2G.4	Random Sampling	106	4	Prediction, Goodness-of-Fit, and Modeling Issues	152
Appendix 2H	Monte Carlo Simulation	106	4.1	Least Squares Prediction	153
2H.1	The Regression Function	106	4.2	Measuring Goodness-of-Fit	156
2H.2	The Random Error	107	4.2.1	Correlation Analysis	158
2H.3	Theoretically True Values	107	4.2.2	Correlation Analysis and R^2	158
2H.4	Creating a Sample of Data	108	4.3	Modeling Issues	160
2H.5	Monte Carlo Objectives	109	4.3.1	The Effects of Scaling the Data	160
2H.6	Monte Carlo Results	109	4.3.2	Choosing a Functional Form	161
2H.7	Random- x Monte Carlo Results	110	4.3.3	A Linear-Log Food Expenditure Model	163
			4.3.4	Using Diagnostic Residual Plots	165
			4.3.5	Are the Regression Errors Normally Distributed?	167
3	Interval Estimation and Hypothesis Testing	112			
3.1	Interval Estimation	113			
3.1.1	The t -Distribution	113			
3.1.2	Obtaining Interval Estimates	115			
3.1.3	The Sampling Context	116			

xiv Table of Contents

4.3.6 Identifying Influential Observations 169

4.4 Polynomial Models 171

4.4.1 Quadratic and Cubic Equations 171

4.5 Log-Linear Models 173

4.5.1 Prediction in the Log-Linear Model 175

4.5.2 A Generalized R^2 Measure 176

4.5.3 Prediction Intervals in the Log-Linear Model 177

4.6 Log-Log Models 177

4.7 Exercises 179

4.7.1 Problems 179

4.7.2 Computer Exercises 185

Appendix 4A Development of a Prediction Interval 192

Appendix 4B The Sum of Squares Decomposition 193

Appendix 4C Mean Squared Error: Estimation and Prediction 193

5 The Multiple Regression Model 196

5.1 Introduction 197

5.1.1 The Economic Model 197

5.1.2 The Econometric Model 198

5.1.3 The General Model 202

5.1.4 Assumptions of the Multiple Regression Model 203

5.2 Estimating the Parameters of the Multiple Regression Model 205

5.2.1 Least Squares Estimation Procedure 205

5.2.2 Estimating the Error Variance σ^2 207

5.2.3 Measuring Goodness-of-Fit 208

5.2.4 Frisch–Waugh–Lovell (FWL) Theorem 209

5.3 Finite Sample Properties of the Least Squares Estimator 211

5.3.1 The Variances and Covariances of the Least Squares Estimators 212

5.3.2 The Distribution of the Least Squares Estimators 214

5.4 Interval Estimation 216

5.4.1 Interval Estimation for a Single Coefficient 216

5.4.2 Interval Estimation for a Linear Combination of Coefficients 217

5.5 Hypothesis Testing 218

5.5.1 Testing the Significance of a Single Coefficient 219

5.5.2 One-Tail Hypothesis Testing for a Single Coefficient 220

5.5.3 Hypothesis Testing for a Linear Combination of Coefficients 221

5.6 Nonlinear Relationships 222

5.7 Large Sample Properties of the Least Squares Estimator 227

5.7.1 Consistency 227

5.7.2 Asymptotic Normality 229

5.7.3 Relaxing Assumptions 230

5.7.4 Inference for a Nonlinear Function of Coefficients 232

5.8 Exercises 234

5.8.1 Problems 234

5.8.2 Computer Exercises 240

Appendix 5A Derivation of Least Squares Estimators 247

Appendix 5B The Delta Method 248

5B.1 Nonlinear Function of a Single Parameter 248

5B.2 Nonlinear Function of Two Parameters 249

Appendix 5C Monte Carlo Simulation 250

5C.1 Least Squares Estimation with Chi-Square Errors 250

5C.2 Monte Carlo Simulation of the Delta Method 252

Appendix 5D Bootstrapping 254

5D.1 Resampling 255

5D.2 Bootstrap Bias Estimate 256

5D.3 Bootstrap Standard Error 256

5D.4 Bootstrap Percentile Interval Estimate 257

5D.5 Asymptotic Refinement 258

6 Further Inference in the Multiple Regression Model 260

6.1 Testing Joint Hypotheses: The F -test 261

6.1.1 Testing the Significance of the Model 264

6.1.2 The Relationship Between t - and F -Tests 265

6.1.3 More General F -Tests 267

6.1.4 Using Computer Software 268

6.1.5 Large Sample Tests 269

6.2 The Use of Nonsample Information 271

6.3 Model Specification 273

6.3.1 Causality versus Prediction 273

6.3.2 Omitted Variables 275

6.3.3 Irrelevant Variables 277

6.3.4	Control Variables	278
6.3.5	Choosing a Model	280
6.3.6	RESET	281
6.4	Prediction	282
6.4.1	Predictive Model Selection Criteria	285
6.5	Poor Data, Collinearity, and Insignificance	288
6.5.1	The Consequences of Collinearity	289
6.5.2	Identifying and Mitigating Collinearity	290
6.5.3	Investigating Influential Observations	293
6.6	Nonlinear Least Squares	294
6.7	Exercises	297
6.7.1	Problems	297
6.7.2	Computer Exercises	303
Appendix 6A	The Statistical Power of F-Tests	311
Appendix 6B	Further Results from the FWL Theorem	315

7 Using Indicator Variables 317

7.1	Indicator Variables	318
7.1.1	Intercept Indicator Variables	318
7.1.2	Slope-Indicator Variables	320
7.2	Applying Indicator Variables	323
7.2.1	Interactions Between Qualitative Factors	323
7.2.2	Qualitative Factors with Several Categories	324
7.2.3	Testing the Equivalence of Two Regressions	326
7.2.4	Controlling for Time	328
7.3	Log-Linear Models	329
7.3.1	A Rough Calculation	330
7.3.2	An Exact Calculation	330
7.4	The Linear Probability Model	331
7.5	Treatment Effects	332
7.5.1	The Difference Estimator	334
7.5.2	Analysis of the Difference Estimator	334
7.5.3	The Differences-in-Differences Estimator	338
7.6	Treatment Effects and Causal Modeling	342
7.6.1	The Nature of Causal Effects	342
7.6.2	Treatment Effect Models	343
7.6.3	Decomposing the Treatment Effect	344
7.6.4	Introducing Control Variables	345
7.6.5	The Overlap Assumption	347
7.6.6	Regression Discontinuity Designs	347

7.7	Exercises	351
7.7.1	Problems	351
7.7.2	Computer Exercises	358
Appendix 7A	Details of Log-Linear Model Interpretation	366
Appendix 7B	Derivation of the Differences-in-Differences Estimator	366
Appendix 7C	The Overlap Assumption: Details	367

8 Heteroskedasticity 368

8.1	The Nature of Heteroskedasticity	369
8.2	Heteroskedasticity in the Multiple Regression Model	370
8.2.1	The Heteroskedastic Regression Model	371
8.2.2	Heteroskedasticity Consequences for the OLS Estimator	373
8.3	Heteroskedasticity Robust Variance Estimator	374
8.4	Generalized Least Squares: Known Form of Variance	375
8.4.1	Transforming the Model: Proportional Heteroskedasticity	375
8.4.2	Weighted Least Squares: Proportional Heteroskedasticity	377
8.5	Generalized Least Squares: Unknown Form of Variance	379
8.5.1	Estimating the Multiplicative Model	381
8.6	Detecting Heteroskedasticity	383
8.6.1	Residual Plots	384
8.6.2	The Goldfeld-Quandt Test	384
8.6.3	A General Test for Conditional Heteroskedasticity	385
8.6.4	The White Test	387
8.6.5	Model Specification and Heteroskedasticity	388
8.7	Heteroskedasticity in the Linear Probability Model	390
8.8	Exercises	391
8.8.1	Problems	391
8.8.2	Computer Exercises	401
Appendix 8A	Properties of the Least Squares Estimator	407
Appendix 8B	Lagrange Multiplier Tests for Heteroskedasticity	408

Appendix 8C Properties of the Least Squares Residuals 410

- 8C.1 Details of Multiplicative Heteroskedasticity Model 411

Appendix 8D Alternative Robust Sandwich Estimators 411

Appendix 8E Monte Carlo Evidence: OLS, GLS, and FGLS 414

9 Regression with Time-Series Data: Stationary Variables 417

9.1 Introduction 418

- 9.1.1 Modeling Dynamic Relationships 420
- 9.1.2 Autocorrelations 424

9.2 Stationarity and Weak Dependence 427

9.3 Forecasting 430

- 9.3.1 Forecast Intervals and Standard Errors 433
- 9.3.2 Assumptions for Forecasting 435
- 9.3.3 Selecting Lag Lengths 436
- 9.3.4 Testing for Granger Causality 437

9.4 Testing for Serially Correlated Errors 438

- 9.4.1 Checking the Correlogram of the Least Squares Residuals 439
- 9.4.2 Lagrange Multiplier Test 440
- 9.4.3 Durbin–Watson Test 443

9.5 Time-Series Regressions for Policy Analysis 443

- 9.5.1 Finite Distributed Lags 445
- 9.5.2 HAC Standard Errors 448
- 9.5.3 Estimation with AR(1) Errors 452
- 9.5.4 Infinite Distributed Lags 456

9.6 Exercises 463

- 9.6.1 Problems 463
- 9.6.2 Computer Exercises 468

Appendix 9A The Durbin–Watson Test 476

- 9A.1 The Durbin–Watson Bounds Test 478

Appendix 9B Properties of an AR(1) Error 479

10 Endogenous Regressors and Moment-Based Estimation 481

10.1 Least Squares Estimation with Endogenous Regressors 482

- 10.1.1 Large Sample Properties of the OLS Estimator 483

- 10.1.2 Why Least Squares Estimation Fails 484

- 10.1.3 Proving the Inconsistency of OLS 486

10.2 Cases in Which x and e are Contemporaneously Correlated 487

- 10.2.1 Measurement Error 487
- 10.2.2 Simultaneous Equations Bias 488
- 10.2.3 Lagged-Dependent Variable Models with Serial Correlation 489
- 10.2.4 Omitted Variables 489

10.3 Estimators Based on the Method of Moments 490

- 10.3.1 Method of Moments Estimation of a Population Mean and Variance 490
- 10.3.2 Method of Moments Estimation in the Simple Regression Model 491
- 10.3.3 Instrumental Variables Estimation in the Simple Regression Model 492
- 10.3.4 The Importance of Using Strong Instruments 493
- 10.3.5 Proving the Consistency of the IV Estimator 494
- 10.3.6 IV Estimation Using Two-Stage Least Squares (2SLS) 495
- 10.3.7 Using Surplus Moment Conditions 496
- 10.3.8 Instrumental Variables Estimation in the Multiple Regression Model 498
- 10.3.9 Assessing Instrument Strength Using the First-Stage Model 500
- 10.3.10 Instrumental Variables Estimation in a General Model 502
- 10.3.11 Additional Issues When Using IV Estimation 504

10.4 Specification Tests 505

- 10.4.1 The Hausman Test for Endogeneity 505
- 10.4.2 The Logic of the Hausman Test 507
- 10.4.3 Testing Instrument Validity 508

10.5 Exercises 510

- 10.5.1 Problems 510
- 10.5.2 Computer Exercises 516

Appendix 10A Testing for Weak Instruments 520

- 10A.1 A Test for Weak Identification 521
- 10A.2 Testing for Weak Identification: Conclusions 525

Appendix 10B Monte Carlo Simulation 525

- 10B.1 Illustrations Using Simulated Data 526
- 10B.2 The Sampling Properties of IV/2SLS 528

11 Simultaneous Equations Models 531

- 11.1 A Supply and Demand Model 532
- 11.2 The Reduced-Form Equations 534
- 11.3 The Failure of Least Squares Estimation 535
 - 11.3.1 Proving the Failure of OLS 535
- 11.4 The Identification Problem 536
- 11.5 Two-Stage Least Squares Estimation 538
 - 11.5.1 The General Two-Stage Least Squares Estimation Procedure 539
 - 11.5.2 The Properties of the Two-Stage Least Squares Estimator 540
- 11.6 Exercises 545
 - 11.6.1 Problems 545
 - 11.6.2 Computer Exercises 551
- Appendix 11A 2SLS Alternatives 557
 - 11A.1 The k -Class of Estimators 557
 - 11A.2 The LIML Estimator 558
 - 11A.3 Monte Carlo Simulation Results 562

12 Regression with Time-Series Data: Nonstationary Variables 563

- 12.1 Stationary and Nonstationary Variables 564
 - 12.1.1 Trend Stationary Variables 567
 - 12.1.2 The First-Order Autoregressive Model 570
 - 12.1.3 Random Walk Models 572
- 12.2 Consequences of Stochastic Trends 574
- 12.3 Unit Root Tests for Stationarity 576
 - 12.3.1 Unit Roots 576
 - 12.3.2 Dickey–Fuller Tests 577
 - 12.3.3 Dickey–Fuller Test with Intercept and No Trend 577
 - 12.3.4 Dickey–Fuller Test with Intercept and Trend 579
 - 12.3.5 Dickey–Fuller Test with No Intercept and No Trend 580
 - 12.3.6 Order of Integration 581
 - 12.3.7 Other Unit Root Tests 582
- 12.4 Cointegration 582
 - 12.4.1 The Error Correction Model 584
- 12.5 Regression When There Is No Cointegration 585
- 12.6 Summary 587

- 12.7 Exercises 588
 - 12.7.1 Problems 588
 - 12.7.2 Computer Exercises 592

13 Vector Error Correction and Vector Autoregressive Models 597

- 13.1 VEC and VAR Models 598
- 13.2 Estimating a Vector Error Correction Model 600
- 13.3 Estimating a VAR Model 601
- 13.4 Impulse Responses and Variance Decompositions 603
 - 13.4.1 Impulse Response Functions 603
 - 13.4.2 Forecast Error Variance Decompositions 605
- 13.5 Exercises 607
 - 13.5.1 Problems 607
 - 13.5.2 Computer Exercises 608

Appendix 13A The Identification Problem 612

14 Time-Varying Volatility and ARCH Models 614

- 14.1 The ARCH Model 615
- 14.2 Time-Varying Volatility 616
- 14.3 Testing, Estimating, and Forecasting 620
- 14.4 Extensions 622
 - 14.4.1 The GARCH Model—Generalized ARCH 622
 - 14.4.2 Allowing for an Asymmetric Effect 623
 - 14.4.3 GARCH-in-Mean and Time-Varying Risk Premium 624
 - 14.4.4 Other Developments 625
- 14.5 Exercises 626
 - 14.5.1 Problems 626
 - 14.5.2 Computer Exercises 627

15 Panel Data Models 634

- 15.1 The Panel Data Regression Function 636
 - 15.1.1 Further Discussion of Unobserved Heterogeneity 638
 - 15.1.2 The Panel Data Regression Exogeneity Assumption 639

xviii Table of Contents

15.1.3 Using OLS to Estimate the Panel Data Regression 639

15.2 The Fixed Effects Estimator 640

15.2.1 The Difference Estimator: $T = 2$ 640

15.2.2 The Within Estimator: $T = 2$ 642

15.2.3 The Within Estimator: $T > 2$ 643

15.2.4 The Least Squares Dummy Variable Model 644

15.3 Panel Data Regression Error Assumptions 646

15.3.1 OLS Estimation with Cluster-Robust Standard Errors 648

15.3.2 Fixed Effects Estimation with Cluster-Robust Standard Errors 650

15.4 The Random Effects Estimator 651

15.4.1 Testing for Random Effects 653

15.4.2 A Hausman Test for Endogeneity in the Random Effects Model 654

15.4.3 A Regression-Based Hausman Test 656

15.4.4 The Hausman–Taylor Estimator 658

15.4.5 Summarizing Panel Data Assumptions 660

15.4.6 Summarizing and Extending Panel Data Model Estimation 661

15.5 Exercises 663

15.5.1 Problems 663

15.5.2 Computer Exercises 670

Appendix 15A Cluster-Robust Standard Errors: Some Details 677

Appendix 15B Estimation of Error Components 679

16 Qualitative and Limited Dependent Variable Models 681

16.1 Introducing Models with Binary Dependent Variables 682

16.1.1 The Linear Probability Model 683

16.2 Modeling Binary Choices 685

16.2.1 The Probit Model for Binary Choice 686

16.2.2 Interpreting the Probit Model 687

16.2.3 Maximum Likelihood Estimation of the Probit Model 690

16.2.4 The Logit Model for Binary Choices 693

16.2.5 Wald Hypothesis Tests 695

16.2.6 Likelihood Ratio Hypothesis Tests 696

16.2.7 Robust Inference in Probit and Logit Models 698

16.2.8 Binary Choice Models with a Continuous Endogenous Variable 698

16.2.9 Binary Choice Models with a Binary Endogenous Variable 699

16.2.10 Binary Endogenous Explanatory Variables 700

16.2.11 Binary Choice Models and Panel Data 701

16.3 Multinomial Logit 702

16.3.1 Multinomial Logit Choice Probabilities 703

16.3.2 Maximum Likelihood Estimation 703

16.3.3 Multinomial Logit Postestimation Analysis 704

16.4 Conditional Logit 707

16.4.1 Conditional Logit Choice Probabilities 707

16.4.2 Conditional Logit Postestimation Analysis 708

16.5 Ordered Choice Models 709

16.5.1 Ordinal Probit Choice Probabilities 710

16.5.2 Ordered Probit Estimation and Interpretation 711

16.6 Models for Count Data 713

16.6.1 Maximum Likelihood Estimation of the Poisson Regression Model 713

16.6.2 Interpreting the Poisson Regression Model 714

16.7 Limited Dependent Variables 717

16.7.1 Maximum Likelihood Estimation of the Simple Linear Regression Model 717

16.7.2 Truncated Regression 718

16.7.3 Censored Samples and Regression 718

16.7.4 Tobit Model Interpretation 720

16.7.5 Sample Selection 723

16.8 Exercises 725

16.8.1 Problems 725

16.8.2 Computer Exercises 733

Appendix 16A Probit Marginal Effects: Details 739

16A.1 Standard Error of Marginal Effect at a Given Point 739

16A.2 Standard Error of Average Marginal Effect 740

Appendix 16B Random Utility Models 741

16B.1 Binary Choice Model 741

16B.2 Probit or Logit? 742

Appendix 16C Using Latent Variables 743

16C.1 Tobit (Tobit Type I) 743

16C.2 Heckit (Tobit Type II) 744

Appendix 16D A Tobit Monte Carlo Experiment 745

A Mathematical Tools 748

A.1 Some Basics 749

- A.1.1 Numbers 749
- A.1.2 Exponents 749
- A.1.3 Scientific Notation 749
- A.1.4 Logarithms and the Number e 750
- A.1.5 Decimals and Percentages 751
- A.1.6 Logarithms and Percentages 751

A.2 Linear Relationships 752

- A.2.1 Slopes and Derivatives 753
- A.2.2 Elasticity 753

A.3 Nonlinear Relationships 753

- A.3.1 Rules for Derivatives 754
- A.3.2 Elasticity of a Nonlinear Relationship 757
- A.3.3 Second Derivatives 757
- A.3.4 Maxima and Minima 758
- A.3.5 Partial Derivatives 759
- A.3.6 Maxima and Minima of Bivariate Functions 760

A.4 Integrals 762

- A.4.1 Computing the Area Under a Curve 762

A.5 Exercises 764

B Probability Concepts 768

B.1 Discrete Random Variables 769

- B.1.1 Expected Value of a Discrete Random Variable 769
- B.1.2 Variance of a Discrete Random Variable 770
- B.1.3 Joint, Marginal, and Conditional Distributions 771
- B.1.4 Expectations Involving Several Random Variables 772
- B.1.5 Covariance and Correlation 773
- B.1.6 Conditional Expectations 774
- B.1.7 Iterated Expectations 774
- B.1.8 Variance Decomposition 774
- B.1.9 Covariance Decomposition 777

B.2 Working with Continuous Random Variables 778

- B.2.1 Probability Calculations 779
- B.2.2 Properties of Continuous Random Variables 780
- B.2.3 Joint, Marginal, and Conditional Probability Distributions 781
- B.2.4 Using Iterated Expectations with Continuous Random Variables 785

- B.2.5 Distributions of Functions of Random Variables 787

- B.2.6 Truncated Random Variables 789

B.3 Some Important Probability Distributions 789

- B.3.1 The Bernoulli Distribution 790
- B.3.2 The Binomial Distribution 790
- B.3.3 The Poisson Distribution 791
- B.3.4 The Uniform Distribution 792
- B.3.5 The Normal Distribution 793
- B.3.6 The Chi-Square Distribution 794
- B.3.7 The t -Distribution 796
- B.3.8 The F -Distribution 797
- B.3.9 The Log-Normal Distribution 799

B.4 Random Numbers 800

- B.4.1 Uniform Random Numbers 805

B.5 Exercises 806

C Review of Statistical Inference 812

C.1 A Sample of Data 813

C.2 An Econometric Model 814

C.3 Estimating the Mean of a Population 815

- C.3.1 The Expected Value of \bar{Y} 816
- C.3.2 The Variance of \bar{Y} 817
- C.3.3 The Sampling Distribution of \bar{Y} 817
- C.3.4 The Central Limit Theorem 818
- C.3.5 Best Linear Unbiased Estimation 820

C.4 Estimating the Population Variance and Other Moments 820

- C.4.1 Estimating the Population Variance 821
- C.4.2 Estimating Higher Moments 821

C.5 Interval Estimation 822

- C.5.1 Interval Estimation: σ^2 Known 822
- C.5.2 Interval Estimation: σ^2 Unknown 825

C.6 Hypothesis Tests About a Population Mean 826

- C.6.1 Components of Hypothesis Tests 826
- C.6.2 One-Tail Tests with Alternative "Greater Than" ($>$) 828
- C.6.3 One-Tail Tests with Alternative "Less Than" ($<$) 829
- C.6.4 Two-Tail Tests with Alternative "Not Equal To" (\neq) 829
- C.6.5 The p -Value 831
- C.6.6 A Comment on Stating Null and Alternative Hypotheses 832

xx Table of Contents

C.6.7	Type I and Type II Errors	833
C.6.8	A Relationship Between Hypothesis Testing and Confidence Intervals	833
C.7	Some Other Useful Tests	834
C.7.1	Testing the Population Variance	834
C.7.2	Testing the Equality of Two Population Means	834
C.7.3	Testing the Ratio of Two Population Variances	835
C.7.4	Testing the Normality of a Population	836
C.8	Introduction to Maximum Likelihood Estimation	837
C.8.1	Inference with Maximum Likelihood Estimators	840
C.8.2	The Variance of the Maximum Likelihood Estimator	841
C.8.3	The Distribution of the Sample Proportion	842
C.8.4	Asymptotic Test Procedures	843
C.9	Algebraic Supplements	848
C.9.1	Derivation of Least Squares Estimator	848
C.9.2	Best Linear Unbiased Estimation	849

C.10	Kernel Density Estimator	851
C.11	Exercises	854
C.11.1	Problems	854
C.11.2	Computer Exercises	857

D Statistical Tables 862

Table D.1	Cumulative Probabilities for the Standard Normal Distribution $\Phi(z) = P(Z \leq z)$	862
Table D.2	Percentiles of the t-distribution	863
Table D.3	Percentiles of the Chi-square Distribution	864
Table D.4	95th Percentile for the F-distribution	865
Table D.5	99th Percentile for the F-distribution	866
Table D.6	Standard Normal pdf Values $\phi(z)$	867

INDEX	869
-------	-----