

Contents

List of Figures	ix
List of Tables	xiii
Preface	xvii
1 Analysis of Economics Data	1
1.1 Statistical Methods	1
1.2 Types of Data	2
1.3 Regression Analysis	4
1.4 Key Concepts	5
1.5 Exercises	6
2 Univariate Data Summary	7
2.1 Summary Statistics for Numerical Data	7
2.2 Charts for Numerical Data	15
2.3 Charts for Numerical Data by Category	19
2.4 Summary and Charts for Categorical Data	22
2.5 Data Transformation	22
2.6 Data Transformations for Time Series Data	24
2.7 Key Concepts	28
2.8 Exercises	29
3 The Sample Mean	35
3.1 Random Variables	35
3.2 Random Samples	38
3.3 Sample Generated by an Experiment: Coin Tosses	39
3.4 Properties of the Sample Mean	41
3.5 Sampling from a Finite Population: 1880 Census	45
3.6 Estimation of the Population Mean	47
3.7 Nonrepresentative Samples	48
3.8 Computer Generation of Random Samples	50
3.9 Key Concepts	52

3.10 Exercises	54
4 Statistical Inference for the Mean	59
4.1 Example: Mean Annual Earnings	59
4.2 t Statistic and t Distribution	61
4.3 Confidence Intervals	65
4.4 Two-Sided Hypothesis Tests	69
4.5 Two-sided Hypothesis Test Examples	73
4.6 One-Sided Directional Hypothesis Tests	76
4.7 Generalization of Confidence Intervals and Hypothesis Tests	79
4.8 Proportions Data	81
4.9 Key Concepts	83
4.10 Exercises	85
5 Bivariate Data Summary	93
5.1 Example: House Price and Size	93
5.2 Two-way Tabulation	95
5.3 Two-way Scatter Plot	96
5.4 Sample Correlation	97
5.5 Regression Line	100
5.6 Measures of Model Fit	105
5.7 Computer Output following OLS Regression	108
5.8 Prediction and Outlying Observations	110
5.9 Regression and Correlation	111
5.10 Causation	112
5.11 Computations for Correlation and Regression	113
5.12 Nonparametric Regression	115
5.13 Key Concepts	117
5.14 Exercises	119
6 The Least Squares Estimator	125
6.1 Population Model for Bivariate Regression	125
6.2 Examples of Sampling from a Population	128
6.3 Properties of the Least Squares Estimator	133
6.4 Estimators of Model Parameters	138
6.5 Key Concepts	140
6.6 Exercises	141
7 Statistical Inference for Bivariate Regression	145
7.1 Example: House Price and Size	145
7.2 t Statistic	146
7.3 Confidence Intervals	148
7.4 Tests of Statistical Significance	150

7.5	Two-Sided Hypothesis Tests	153
7.6	One-Sided Directional Hypothesis Tests	156
7.7	Robust Standard Errors	158
7.8	Key Concepts	160
7.9	Exercises	161
8	Case Studies for Bivariate Regression	165
8.1	Health Outcomes across Countries	165
8.2	Health Expenditures across Countries	168
8.3	Capital Asset Pricing Model	170
8.4	Output and Unemployment in the U.S.	173
8.5	Exercises	176
9	Models with Natural Logarithms	179
9.1	Natural Logarithm Function	179
9.2	Semi-elasticities and Elasticities	181
9.3	Log-linear, Log-log and Linear-log models	182
9.4	Example: Earnings and Education	185
9.5	Further Uses of the Natural Logarithm	187
9.6	Exponential Function	190
9.7	Key Concepts	191
9.8	Exercises	192
10	Data Summary with Multiple Regression	197
10.1	Example: House Price and Characteristics	197
10.2	Two-way Scatter Plots	199
10.3	Correlation	200
10.4	Regression Line	201
10.5	Interpretation of Slope Coefficients	204
10.6	Model Fit	206
10.7	Computer Output Following Multiple Regression	209
10.8	Inestimable Models	210
10.9	Key Concepts	211
10.10	Exercises	213
11	Statistical Inference for Multiple Regression	217
11.1	Properties of the Least Squares Estimator	217
11.2	Estimators of Model Parameters	221
11.3	Confidence Intervals	222
11.4	Hypothesis Tests on a Single Parameter	223
11.5	Joint Hypothesis Tests	227
11.6	F Statistic under Assumptions 1-4	232
11.7	Presentation of Regression Results	235

11.8 Key Concepts	236
11.9 Exercises	237
12 Further Topics in Multiple Regression	241
12.1 Inference with Robust Standard Errors	241
12.2 Prediction	252
12.3 Nonrepresentative Samples	258
12.4 Best Estimation	260
12.5 Best Confidence Intervals	261
12.6 Best Hypothesis Tests	262
12.7 Data Science and Big Data: An Overview	264
12.8 Bayesian Methods: An Overview	265
12.9 A Brief History of Statistics, Regression and Econometrics	266
12.10 Key Concepts	267
12.11 Exercises	268
13 Case Studies for Multiple Regression	275
13.1 School Academic Performance	275
13.2 Cobb-Douglas Production Function	280
13.3 Phillips Curve	284
13.4 Automobile Fuel Efficiency	288
13.5 Rand Health Insurance Experiment	291
13.6 Access to Health Care and Health Outcomes	295
13.7 Gains from Political Incumbency	298
13.8 Institutions and Country GDP	301
13.9 From Raw Data to Final Data	303
13.10 Exercises	308
14 Regression with Indicator Variables	315
14.1 Example: Earnings, Gender, Education and Type of Worker	315
14.2 Regression on just a Single Indicator Variable	315
14.3 Regression on an Indicator Variable and Additional Regressors	319
14.4 Regression with Sets of Indicator Variables	322
14.5 Key Concepts	328
14.6 Exercises	329
15 Regression with Transformed Variables	333
15.1 Example: Earnings, Gender, Education and Type of Worker	333
15.2 Marginal Effects for Nonlinear Models	333
15.3 Quadratic Model and Polynomial Models	337
15.4 Interacted Regressors	340
15.5 Log-linear and Log-log Models	342
15.6 Prediction from Log-Linear and Log-Log Models	344

15.7 Models with a Mix of Regressor Types	346
15.8 Key Concepts	348
15.9 Exercises	349
16 Checking the Model and Data	355
16.1 Multicollinear Data	356
16.2 Model Assumptions Revisited	358
16.3 Incorrect Population Model	359
16.4 Regressors Correlated with Errors	362
16.5 Heteroskedastic Errors	363
16.6 Correlated Errors	364
16.7 Example: Democracy and Growth	365
16.8 Diagnostics	367
16.9 Key Concepts	371
16.10 Exercises	372
17 Special Topics	377
17.1 Cross-Section Data	377
17.2 Panel Data	380
17.3 Panel Data Example: NBA Team Revenue	384
17.4 Instrumental Variables	387
17.5 Causal Inference: An Overview	391
17.6 Nonlinear Regression Models	396
17.7 Time-Series Data	402
17.8 Time Series Example: U.S. Treasury Security Interest Rates	409
17.9 Further Reading	416
17.10 Key Concepts	416
17.11 Exercises	417
A Using Statistical Packages	429
A.1 General Issues	429
A.2 Stata Essentials	434
A.3 R Essentials	436
A.4 Gretl Essentials	440
A.5 Eviews Essentials	445
A.6 Excel and Google Sheets Spreadsheet Applications	449
A.7 Critical values and p-values	451
B Some Essentials of Probability Theory	453
B.1 Probability Theory for a Single Random Variable	453
B.2 Probability Theory for the Sample Mean	457
B.3 Probability Theory for Two Related Random Variables	459

C	Properties of OLS, IV and ML Estimators	463
C.1	OLS with Independent Homoskedastic Errors	463
C.2	Robust Standard Errors	466
C.3	Instrumental Variables Estimation	470
C.4	OLS with Matrix Algebra	471
C.5	Maximum Likelihood Estimation	473
C.6	Exercises	474
D	Solutions to Selected Exercises	475
D.1	Solutions: Analysis of Economics Data	475
D.2	Solutions: Univariate Data Summary	475
D.3	Solutions: The Sample Mean	476
D.4	Solutions: Statistical Inference for the Mean	477
D.5	Solutions: Bivariate Data Summary	478
D.6	Solutions: The Least Squares Estimator	479
D.7	Solutions: Statistical Inference for Bivariate Regression	480
D.8	Solutions: Case Studies for Bivariate Regression	481
D.9	Solutions: Models with Natural Logarithm	481
D.10	Solutions: Data Summary with Multiple Regression	482
D.11	Solutions: Statistical Inference for Multiple Regression	483
D.12	Solutions: Further Topics in Multiple Regression	484
D.13	Solutions: Case Studies for Multiple Regression	485
D.14	Solutions: Models with Indicator Variables	486
D.15	Solutions: Models with Transformed Variables	486
D.16	Solutions: Checking the Model and Data	487
D.17	Solutions: Special Topics	489
E	Tables for Key Distributions	493
F	References	501
	Index	503