

Contents

Preface	xxv
Chapter 1: Introduction	1
1.1 Multiple Regression/Correlation as a General Data-Analytic System	1
1.1.1 Overview	1
1.1.2 Testing Hypotheses Using Multiple Regression/Correlation: Some Examples	2
1.1.3 Multiple Regression/Correlation in Prediction Models	3
1.2 A Comparison of Multiple Regression/Correlation and Analysis of Variance Approaches	4
1.2.1 Historical Background	4
1.2.2 Hypothesis Testing and Effect Sizes	5
1.3 Multiple Regression/Correlation and the Complexity of Behavioral Science	6
1.3.1 Multiplicity of Influences	6
1.3.2 Correlation Among Research Factors and Partialing	6
1.3.3 Form of Information	7
1.3.4 Shape of Relationship	8
1.3.5 General and Conditional Relationships	9
1.4 Orientation of the Book	10
1.4.1 Nonmathematical	11
1.4.2 Applied	11
1.4.3 Data-Analytic	12
1.4.4 Inference Orientation and Specification Error	13
1.5 Computation, the Computer, and Numerical Results	14
1.5.1 Computation	14

1.5.2	Numerical Results: Reporting and Rounding	14
1.5.3	Significance Tests, Confidence Intervals, and Appendix Tables	15
1.6	The Spectrum of Behavioral Science	16
1.7	Plan for the Book	16
1.7.1	Content	16
1.7.2	Structure: Numbering of Sections, Tables, and Equations	17
1.8	Summary	18

Chapter 2: Bivariate Correlation and Regression

19

2.1	Tabular and Graphic Representations of Relationships	19
2.2	The Index of Linear Correlation Between Two Variables: The Pearson Product Moment Correlation Coefficient	23
2.2.1	Standard Scores: Making Units Comparable	23
2.2.2	The Product Moment Correlation as a Function of Differences Between z Scores	26
2.3	Alternative Formulas for the Product Moment Correlation Coefficient	28
2.3.1	r as the Average Product of z Scores	28
2.3.2	Raw Score Formulas for r	29
2.3.3	Point Biserial r	29
2.3.4	Phi (ϕ) Coefficient	30
2.3.5	Rank Correlation	31
2.4	Regression Coefficients: Estimating Y From X	32
2.5	Regression Toward the Mean	36
2.6	The Standard Error of Estimate and Measures of the Strength of Association	37
2.7	Summary of Definitions and Interpretations	41
2.8	Statistical Inference With Regression and Correlation Coefficients	41
2.8.1	Assumptions Underlying Statistical Inference With B_{YX} , B_0 , \hat{Y}_i , and r_{XY}	41
2.8.2	Estimation With Confidence Intervals	42
2.8.3	Null Hypothesis Significance Tests (NHSTs)	47
2.8.4	Confidence Limits and Null Hypothesis Significance Testing	50
2.9	Precision and Power	50
2.9.1	Precision of Estimation	50
2.9.2	Power of Null Hypothesis Significance Tests	51
2.10	Factors Affecting the Size of r	53
2.10.1	The Distributions of X and Y	53
2.10.2	The Reliability of the Variables	55
2.10.3	Restriction of Range	57
2.10.4	Part-Whole Correlations	59
2.10.5	Ratio or Index Variables	60
2.10.6	Curvilinear Relationships	62
2.11	Summary	62

Chapter 3: Multiple Regression/Correlation With Two or More Independent Variables

64

3.1	Introduction: Regression and Causal Models	64
3.1.1	What Is a Cause?	64
3.1.2	Diagrammatic Representation of Causal Models	65
3.2	Regression With Two Independent Variables	66
3.3	Measures of Association With Two Independent Variables	69
3.3.1	Multiple R and R^2	69
3.3.2	Semipartial Correlation Coefficients and Increments to R^2	72
3.3.3	Partial Correlation Coefficients	74
3.4	Patterns of Association Between Y and Two Independent Variables	75
3.4.1	Direct and Indirect Effects	75
3.4.2	Partial Redundancy	76
3.4.3	Suppression in Regression Models	77
3.4.4	Spurious Effects and Entirely Indirect Effects	78
3.5	Multiple Regression/Correlation With k Independent Variables	79
3.5.1	Introduction: Components of the Prediction Equation	79
3.5.2	Partial Regression Coefficients	80
3.5.3	R , R^2 , and Shrunk R^2	82
3.5.4	sr and sr^2	84
3.5.5	pr and pr^2	85
3.5.6	Example of Interpretation of Partial Coefficients	85
3.6	Statistical Inference With k Independent Variables	86
3.6.1	Standard Errors and Confidence Intervals for B and β	86
3.6.2	Confidence Intervals for R^2	88
3.6.3	Confidence Intervals for Differences Between Independent R^2 s	88
3.6.4	Statistical Tests on Multiple and Partial Coefficients	88
3.7	Statistical Precision and Power Analysis	90
3.7.1	Introduction: Research Goals and the Null Hypothesis	90
3.7.2	The Precision and Power of R^2	91
3.7.3	Precision and Power Analysis for Partial Coefficients	93
3.8	Using Multiple Regression Equations in Prediction	95
3.8.1	Prediction of Y for a New Observation	95
3.8.2	Correlation of Individual Variables With Predicted Values	96
3.8.3	Cross-Validation and Unit Weighting	97
3.8.4	Multicollinearity	98
3.9	Summary	99

**Chapter 4: Data Visualization, Exploration, and Assumption
Checking: Diagnosing and Solving
Regression Problems I**

101

4.1	Introduction	101
4.2	Some Useful Graphical Displays of the Original Data	102
4.2.1	Univariate Displays	103
4.2.2	Bivariate Displays	110
4.2.3	Correlation and Scatterplot Matrices	115
4.3	Assumptions and Ordinary Least Squares Regression	117
4.3.1	Assumptions Underlying Multiple Linear Regression	117
4.3.2	Ordinary Least Squares Estimation	124
4.4	Detecting Violations of Assumptions	125
4.4.1	Form of the Relationship	125
4.4.2	Omitted Independent Variables	127
4.4.3	Measurement Error	129
4.4.4	Homoscedasticity of Residuals	130
4.4.5	Nonindependence of Residuals	134
4.4.6	Normality of Residuals	137
4.5	Remedies: Alternative Approaches When Problems Are Detected	141
4.5.1	Form of the Relationship	141
4.5.2	Inclusion of All Relevant Independent Variables	143
4.5.3	Measurement Error in the Independent Variables	144
4.5.4	Nonconstant Variance	145
4.5.5	Nonindependence of Residuals	147
4.6	Summary	150

**Chapter 5: Data-Analytic Strategies Using Multiple
Regression/Correlation**

151

5.1	Research Questions Answered by Correlations and Their Squares	151
5.1.1	Net Contribution to Prediction	152
5.1.2	Indices of Differential Validity	152
5.1.3	Comparisons of Predictive Utility	152
5.1.4	Attribution of a Fraction of the XY Relationship to a Third Variable	153
5.1.5	Which of Two Variables Accounts for More of the XY Relationship?	153
5.1.6	Are the Various Squared Correlations in One Population Different From Those in Another Given the Same Variables?	154
5.2	Research Questions Answered by B Or β	154
5.2.1	Regression Coefficients as Reflections of Causal Effects	154
5.2.2	Alternative Approaches to Making B_{YX} Substantively Meaningful	154
5.2.3	Are the Effects of a Set of Independent Variables on Two Different Outcomes in a Sample Different?	157

5.2.4	What Are the Reciprocal Effects of Two Variables on One Another?	157
5.3	Hierarchical Analysis Variables in Multiple Regression/Correlation	158
5.3.1	Causal Priority and the Removal of Confounding Variables	158
5.3.2	Research Relevance	160
5.3.3	Examination of Alternative Hierarchical Sequences of Independent Variable Sets	160
5.3.4	Stepwise Regression	161
5.4	The Analysis of Sets of Independent Variables	162
5.4.1	Types of Sets	162
5.4.2	The Simultaneous and Hierarchical Analyses of Sets	164
5.4.3	Variance Proportions for Sets and the Ballantine Again	166
5.4.4	B and β Coefficients for Variables Within Sets	169
5.5	Significance Testing for Sets	171
5.5.1	Application in Hierarchical Analysis	172
5.5.2	Application in Simultaneous Analysis	173
5.5.3	Using Computer Output to Determine Statistical Significance	174
5.5.4	An Alternative F Test: Using Model 2 Error Estimate From the Final Model	174
5.6	Power Analysis for Sets	176
5.6.1	Determining n^* for the F Test of sR_B^2 with Model 1 or Model 2 Error	177
5.6.2	Estimating the Population sR^2 Values	179
5.6.3	Setting Power for n^*	180
5.6.4	Reconciling Different n^* s	180
5.6.5	Power as a Function of n	181
5.6.6	Tactics of Power Analysis	182
5.7	Statistical Inference Strategy in Multiple Regression/Correlation	182
5.7.1	Controlling and Balancing Type I and Type II Errors in Inference	182
5.7.2	Less Is More	185
5.7.3	Least Is Last	186
5.7.4	Adaptation of Fisher's Protected t Test	187
5.7.5	Statistical Inference and the Stage of Scientific Investigations	190
5.8	Summary	190

Chapter 6: Quantitative Scales, Curvilinear Relationships, and Transformations

193

6.1	Introduction	193
6.1.1	What Do We Mean by Linear Regression?	193

6.1.2	Linearity in the Variables and Linear Multiple Regression	194
6.1.3	Four Approaches to Examining Nonlinear Relationships in Multiple Regression	195
6.2	Power Polynomials	196
6.2.1	Method	196
6.2.2	An Example: Quadratic Fit	198
6.2.3	Centering Predictors in Polynomial Equations	201
6.2.4	Relationship of Test of Significance of Highest Order Coefficient and Gain in Prediction	204
6.2.5	Interpreting Polynomial Regression Results	205
6.2.6	Another Example: A Cubic Fit	207
6.2.7	Strategy and Limitations	209
6.2.8	More Complex Equations	213
6.3	Orthogonal Polynomials	214
6.3.1	The Cubic Example Revisited	216
6.3.2	Unequal n and Unequal Intervals	219
6.3.3	Applications and Discussion	220
6.4	Nonlinear Transformations	221
6.4.1	Purposes of Transformation and the Nature of Transformations	221
6.4.2	The Conceptual Basis of Transformations and Model Checking Before and After Transformation—Is It Always Ideal to Transform?	223
6.4.3	Logarithms and Exponents; Additive and Proportional Relationships	223
6.4.4	Linearizing Relationships	225
6.4.5	Linearizing Relationships Based on Strong Theoretical Models	227
6.4.6	Linearizing Relationships Based on Weak Theoretical Models	232
6.4.7	Empirically Driven Transformations in the Absence of Strong or Weak Models	233
6.4.8	Empirically Driven Transformation for Linearization: The Ladder of Re-expression and the Bulging Rule	233
6.4.9	Empirically Driven Transformation for Linearization in the Absence of Models: Box-Cox Family of Power Transformations on Y	236
6.4.10	Empirically Driven Transformation for Linearization in the Absence of Models: Box-Tidwell Family of Power Transformations on X	239
6.4.11	Linearization of Relationships With Correlations: Fisher z' Transform of r	240
6.4.12	Transformations That Linearize Relationships for Counts and Proportions	240
6.4.13	Variance Stabilizing Transformations and Alternatives for Treatment of Heteroscedasticity	244
6.4.14	Transformations to Normalize Variables	246
6.4.15	Diagnostics Following Transformation	247

6.4.16	Measuring and Comparing Model Fit	248
6.4.17	Second-Order Polynomial Numerical Example Revisited	248
6.4.18	When to Transform and the Choice of Transformation	249
6.5	Nonlinear Regression	251
6.6	Nonparametric Regression	252
6.7	Summary	253

Chapter 7: Interactions Among Continuous Variables **255**

7.1	Introduction	255
7.1.1	Interactions Versus Additive Effects	256
7.1.2	Conditional First-Order Effects in Equations Containing Interactions	259
7.2	Centering Predictors and the Interpretation of Regression Coefficients in Equations Containing Interactions	261
7.2.1	Regression with Centered Predictors	261
7.2.2	Relationship Between Regression Coefficients in the Uncentered and Centered Equations	262
7.2.3	Centered Equations With No Interaction	262
7.2.4	Essential Versus Nonessential Multicollinearity	264
7.2.5	Centered Equations With Interactions	264
7.2.6	The Highest Order Interaction in the Centered Versus Uncentered Equation	266
7.2.7	Do Not Center Y	266
7.2.8	A Recommendation for Centering	266
7.3	Simple Regression Equations and Simple Slopes	267
7.3.1	Plotting Interactions	269
7.3.2	Moderator Variables	269
7.3.3	Simple Regression Equations	269
7.3.4	Overall Regression Coefficient and Simple Slope at the Mean	270
7.3.5	Simple Slopes From Uncentered Versus Centered Equations Are Identical	271
7.3.6	Linear by Linear Interactions	271
7.3.7	Interpreting Interactions in Multiple Regression and Analysis of Variance	272
7.4	Post Hoc Probing of Interactions	272
7.4.1	Standard Error of Simple Slopes	272
7.4.2	Equation Dependence of Simple Slopes and Their Standard Errors	273
7.4.3	Tests of Significance of Simple Slopes	273
7.4.4	Confidence Intervals Around Simple Slopes	274
7.4.5	A Numerical Example	275
7.4.6	The Uncentered Regression Equation Revisited	281
7.4.7	First-Order Coefficients in Equations Without and With Interactions	281
7.4.8	Interpretation and the Range of Data	282

7.5	Standardized Estimates for Equations Containing Interactions	282
7.6	Interactions as Partialled Effects: Building Regression Equations With Interactions	284
7.7	Patterns of First-Order and Interactive Effects	285
7.7.1	Three Theoretically Meaningful Patterns of First-Order and Interaction Effects	285
7.7.2	Ordinal Versus Disordinal Interactions	286
7.8	Three-Predictor Interactions in Multiple Regression	290
7.9	Curvilinear by Linear Interactions	292
7.10	Interactions Among Sets of Variables	295
7.11	Issues in the Detection of Interactions: Reliability, Predictor Distributions, Model Specification	297
7.11.1	Variable Reliability and Power to Detect Interactions	297
7.11.2	Sampling Designs to Enhance Power to Detect Interactions—Optimal Design	298
7.11.3	Difficulty in Distinguishing Interactions Versus Curvilinear Effects	299
7.12	Summary	300
 Chapter 8: Categorical or Nominal Independent Variables		302
8.1	Introduction	302
8.1.1	Categories as a Set of Independent Variables	302
8.1.2	The Representation of Categories or Nominal Scales	302
8.2	Dummy-Variable Coding	303
8.2.1	Coding the Groups	303
8.2.2	Pearson Correlations of Dummy Variables With <i>Y</i>	308
8.2.3	Correlations Among Dummy-Coded Variables	311
8.2.4	Multiple Correlation of the Dummy-Variable Set With <i>Y</i>	311
8.2.5	Regression Coefficients for Dummy Variables	312
8.2.6	Partial and Semipartial Correlations for Dummy Variables	316
8.2.7	Dummy-Variable Multiple Regression/Correlation and One-Way Analysis of Variance	317
8.2.8	A Cautionary Note: Dummy-Variable-Like Coding Systems	319
8.2.9	Dummy-Variable Coding When Groups Are Not Mutually Exclusive	320
8.3	Unweighted Effects Coding	320
8.3.1	Introduction: Unweighted and Weighted Effects Coding	320
8.3.2	Constructing Unweighted Effects Codes	321
8.3.3	The R^2 and the r_{Yi} s for Unweighted Effects Codes	324
8.3.4	Regression Coefficients and Other Partial Effects in Unweighted Code Sets	325

8.4	Weighted Effects Coding	328
8.4.1	Selection Considerations for Weighted Effects Coding	328
8.4.2	Constructing Weighted Effects	328
8.4.3	The R^2 and \bar{R}^2 for Weighted Effects Codes	330
8.4.4	Interpretation and Testing of B With Unweighted Codes	331
8.5	Contrast Coding	332
8.5.1	Considerations in the Selection of a Contrast Coding Scheme	332
8.5.2	Constructing Contrast Codes	333
8.5.3	The R^2 and \bar{R}^2	337
8.5.4	Partial Regression Coefficients	337
8.5.5	Statistical Power and the Choice of Contrast Codes	340
8.6	Nonsense Coding	341
8.7	Coding Schemes in the Context of Other Independent Variables	342
8.7.1	Combining Nominal and Continuous Independent Variables	342
8.7.2	Calculating Adjusted Means for Nominal Independent Variables	343
8.7.3	Adjusted Means for Combinations of Nominal and Quantitative Independent Variables	344
8.7.4	Adjusted Means for More Than Two Groups and Alternative Coding Methods	348
8.7.5	Multiple Regression/Correlation With Nominal Independent Variables and the Analysis of Covariance	350
8.8	Summary	351

Chapter 9: Interactions With Categorical Variables

354

9.1	Nominal Scale by Nominal Scale Interactions	354
9.1.1	The 2 by 2 Design	354
9.1.2	Regression Analyses of Multiple Sets of Nominal Variables With More Than Two Categories	361
9.2	Interactions Involving More Than Two Nominal Scales	366
9.2.1	An Example of Three Nominal Scales Coded by Alternative Methods	367
9.2.2	Interactions Among Nominal Scales in Which Not All Combinations Are Considered	372
9.2.3	What If the Categories for One or More Nominal “Scales” Are Not Mutually Exclusive?	373
9.2.4	Consideration of pr , β , and Variance Proportions for Nominal Scale Interaction Variables	374
9.2.5	Summary of Issues and Recommendations for Interactions Among Nominal Scales	374
9.3	Nominal Scale by Continuous Variable Interactions	375
9.3.1	A Reminder on Centering	375

9.3.2	Interactions of a Continuous Variable With Dummy-Variable Coded Groups	375
9.3.3	Interactions Using Weighted or Unweighted Effects Codes	378
9.3.4	Interactions With a Contrast-Coded Nominal Scale	379
9.3.5	Interactions Coded to Estimate Simple Slopes of Groups	380
9.3.6	Categorical Variable Interactions With Nonlinear Effects of Scaled Independent Variables	383
9.3.7	Interactions of a Scale With Two or More Categorical Variables	386
9.4	Summary	388

Chapter 10: Outliers and Multicollinearity: Diagnosing and Solving Regression Problems II

390

10.1	Introduction	390
10.2	Outliers: Introduction and Illustration	391
10.3	Detecting Outliers: Regression Diagnostics	394
10.3.1	Extremity on the Independent Variables: Leverage	394
10.3.2	Extremity on Y: Discrepancy	398
10.3.3	Influence on the Regression Estimates	402
10.3.4	Location of Outlying Points and Diagnostic Statistics	406
10.3.5	Summary and Suggestions	409
10.4	Sources of Outliers and Possible Remedial Actions	411
10.4.1	Sources of Outliers	411
10.4.2	Remedial Actions	415
10.5	Multicollinearity	419
10.5.1	Exact Collinearity	419
10.5.2	Multicollinearity: A Numerical Illustration	420
10.5.3	Measures of the Degree of Multicollinearity	422
10.6	Remedies for Multicollinearity	425
10.6.1	Model Respecification	426
10.6.2	Collection of Additional Data	427
10.6.3	Ridge Regression	427
10.6.4	Principal Components Regression	428
10.6.5	Summary of Multicollinearity Considerations	429
10.7	Summary	430

Chapter 11: Missing Data

431

11.1	Basic Issues in Handling Missing Data	431
11.1.1	Minimize Missing Data	431
11.1.2	Types of Missing Data	432
11.1.3	Traditional Approaches to Missing Data	433

11.2	Missing Data in Nominal Scales	435
11.2.1	Coding Nominal Scale X for Missing Data	435
11.2.2	Missing Data on Two Dichotomies	439
11.2.3	Estimation Using the EM Algorithm	440
11.3	Missing Data in Quantitative Scales	442
11.3.1	Available Alternatives	442
11.3.2	Imputation of Values for Missing Cases	444
11.3.3	Modeling Solutions to Missing Data in Scaled Variables	447
11.3.4	An Illustrative Comparison of Alternative Methods	447
11.3.5	Rules of Thumb	450
11.4	Summary	450

Chapter 12: Multiple Regression/Correlation and Causal Models

452

12.1	Introduction	452
12.1.1	Limits on the Current Discussion and the Relationship Between Causal Analysis and Analysis of Covariance	452
12.1.2	Theories and Multiple Regression/Correlation Models That Estimate and Test Them	454
12.1.3	Kinds of Variables in Causal Models	457
12.1.4	Regression Models as Causal Models	459
12.2	Models Without Reciprocal Causation	460
12.2.1	Direct and Indirect Effects	460
12.2.2	Path Analysis and Path Coefficients	464
12.2.3	Hierarchical Analysis and Reduced Form Equations	465
12.2.4	Partial Causal Models and the Hierarchical Analysis of Sets	466
12.2.5	Testing Model Elements	467
12.3	Models With Reciprocal Causation	467
12.4	Identification and Overidentification	468
12.4.1	Just Identified Models	468
12.4.2	Overidentification	468
12.4.3	Underidentification	469
12.5	Latent Variable Models	469
12.5.1	An Example of a Latent Variable Model	469
12.5.2	How Latent Variables Are Estimated	471
12.5.3	Fixed and Free Estimates in Latent Variable Models	472
12.5.4	Goodness-of-Fit Tests of Latent Variable Models	472
12.5.5	Latent Variable Models and the Correction for Attenuation	473
12.5.6	Characteristics of Data Sets That Make Latent Variable Analysis the Method of Choice	474
12.6	A Review of Causal Model and Statistical Assumptions	475

12.6.1	Specification Error	475
12.6.2	Identification Error	475
12.7	Comparisons of Causal Models	476
12.7.1	Nested Models	476
12.7.2	Longitudinal Data in Causal Models	476
12.8	Summary	477

Chapter 13: Alternative Regression Models: Logistic, Poisson Regression, and the Generalized Linear Model **479**

13.1	Ordinary Least Squares Regression Revisited	479
13.1.1	Three Characteristics of Ordinary Least Squares Regression	480
13.1.2	The Generalized Linear Model	480
13.1.3	Relationship of Dichotomous and Count Dependent Variables Y to a Predictor	481
13.2	Dichotomous Outcomes and Logistic Regression	482
13.2.1	Extending Linear Regression: The Linear Probability Model and Discriminant Analysis	483
13.2.2	The Nonlinear Transformation From Predictor to Predicted Scores: Probit and Logistic Transformation	485
13.2.3	The Logistic Regression Equation	486
13.2.4	Numerical Example: Three Forms of the Logistic Regression Equation	487
13.2.5	Understanding the Coefficients for the Predictor in Logistic Regression	492
13.2.6	Multiple Logistic Regression	493
13.2.7	Numerical Example	494
13.2.8	Confidence Intervals on Regression Coefficients and Odds Ratios	497
13.2.9	Estimation of the Regression Model: Maximum Likelihood	498
13.2.10	Deviances: Indices of Overall Fit of the Logistic Regression Model	499
13.2.11	Multiple R^2 Analogs in Logistic Regression	502
13.2.12	Testing Significance of Overall Model Fit: The Likelihood Ratio Test and the Test of Model Deviance	504
13.2.13	χ^2 Test for the Significance of a Single Predictor in a Multiple Logistic Regression Equation	507
13.2.14	Hierarchical Logistic Regression: Likelihood Ratio χ^2 Test for the Significance of a Set of Predictors Above and Beyond Another Set	508
13.2.15	Akaike's Information Criterion and the Bayesian Information Criterion for Model Comparison	509
13.2.16	Some Treachery in Variable Scaling and Interpretation of the Odds Ratio	509

13.2.17	Regression Diagnostics in Logistic Regression	512
13.2.18	Sparseness of Data	516
13.2.19	Classification of Cases	516
13.3	Extensions of Logistic Regression to Multiple Response Categories: Polytomous Logistic Regression and Ordinal Logistic Regression	519
13.3.1	Polytomous Logistic Regression	519
13.3.2	Nested Dichotomies	520
13.3.3	Ordinal Logistic Regression	522
13.4	Models for Count Data: Poisson Regression and Alternatives	525
13.4.1	Linear Regression Applied to Count Data	525
13.4.2	Poisson Probability Distribution	526
13.4.3	Poisson Regression Analysis	528
13.4.4	Overdispersion and Alternative Models	530
13.4.5	Independence of Observations	532
13.4.6	Sources on Poisson Regression	532
13.5	Full Circle: Parallels Between Logistic and Poisson Regression, and the Generalized Linear Model	532
13.5.1	Parallels Between Poisson and Logistic Regression	532
13.5.2	The Generalized Linear Model Revisited	534
13.6	Summary	535

Chapter 14: Random Coefficient Regression and Multilevel Models

536

14.1	Clustering Within Data Sets	536
14.1.1	Clustering, Alpha Inflation, and the Intraclass Correlation	537
14.1.2	Estimating the Intraclass Correlation	538
14.2	Analysis of Clustered Data With Ordinary Least Squares Approaches	539
14.2.1	Numerical Example, Analysis of Clustered Data With Ordinary Least Squares Regression	541
14.3	The Random Coefficient Regression Model	543
14.4	Random Coefficient Regression Model and Multilevel Data Structure	544
14.4.1	Ordinary Least Squares (Fixed Effects) Regression Revisited	544
14.4.2	Fixed and Random Variables	544
14.4.3	Clustering and Hierarchically Structured Data	545
14.4.4	Structure of the Random Coefficient Regression Model	545
14.4.5	Level 1 Equations	546
14.4.6	Level 2 Equations	547
14.4.7	Mixed Model Equation for Random Coefficient Regression	548
14.4.8	Variance Components—New Parameters in the Multilevel Model	548
14.4.9	Variance Components and Random Coefficient Versus Ordinary Least Squares (Fixed Effects) Regression	549

14.4.10	Parameters of the Random Coefficient Regression Model: Fixed and Random Effects	550
14.5	Numerical Example: Analysis of Clustered Data With Random Coefficient Regression	550
14.5.1	Unconditional Cell Means Model and the Intraclass Correlation	551
14.5.2	Testing the Fixed and Random Parts of the Random Coefficient Regression Model	552
14.6	Clustering as a Meaningful Aspect of the Data	553
14.7	Multilevel Modeling With a Predictor at Level 2	553
14.7.1	Level 1 Equations	553
14.7.2	Revised Level 2 Equations	554
14.7.3	Mixed Model Equation With Level 1 Predictor and Level 2 Predictor of Intercept and Slope and the Cross-Level Interaction	554
14.8	An Experimental Design as a Multilevel Data Structure: Combining Experimental Manipulation With Individual Differences	555
14.9	Numerical Example: Multilevel Analysis	556
14.10	Estimation of the Multilevel Model Parameters: Fixed Effects, Variance Components, and Level 1 Equations	560
14.10.1	Fixed Effects and Variance Components	560
14.10.2	An Equation for Each Group: Empirical Bayes Estimates of Level 1 Coefficients	560
14.11	Statistical Tests in Multilevel Models	563
14.11.1	Fixed Effects	563
14.11.2	Variance Components	563
14.12	Some Model Specification Issues	564
14.12.1	The Same Variable at Two Levels	564
14.12.2	Centering in Multilevel Models	564
14.13	Statistical Power of Multilevel Models	565
14.14	Choosing Between the Fixed Effects Model and the Random Coefficient Model	565
14.15	Sources on Multilevel Modeling	566
14.16	Multilevel Models Applied to Repeated Measures Data	566
14.17	Summary	567

Chapter 15: Longitudinal Regression Methods

568

15.1	Introduction	568
15.1.1	Chapter Goals	568
15.1.2	Purposes of Gathering Data on Multiple Occasions	569
15.2	Analyses of Two-Time-Point Data	569
15.2.1	Change or Regressed Change?	570
15.2.2	Alternative Regression Models for Effects Over a Single Unit of Time	571
15.2.3	Three- or Four-Time-Point Data	573
15.3	Repeated Measure Analysis of Variance	573

15.3.1	Multiple Error Terms in Repeated Measure Analysis of Variance	574
15.3.2	Trend Analysis in Analysis of Variance	575
15.3.3	Repeated Measure Analysis of Variance in Which Time Is Not the Issue	576
15.4	Multilevel Regression of Individual Changes Over Time	578
15.4.1	Patterns of Individual Change Over Time	578
15.4.2	Adding Other Fixed Predictors to the Model	582
15.4.3	Individual Differences in Variation Around Individual Slopes	583
15.4.4	Alternative Developmental Models and Error Structures	584
15.4.5	Alternative Link Functions for Predicting Y From Time	586
15.4.6	Unbalanced Data: Variable Timing and Missing Data	587
15.5	Latent Growth Models: Structural Equation Model Representation of Multilevel Data	588
15.5.1	Estimation of Changes in True Scores	589
15.5.2	Representation of Latent Growth Models in Structural Equation Model Diagrams	589
15.5.3	Comparison of Multilevel Regression and Structural Equation Model Analysis of Change	594
15.6	Time Varying Independent Variables	595
15.7	Survival Analysis	596
15.7.1	Regression Analysis of Time Until Outcome and the Problem of Censoring	596
15.7.2	Extension to Time-Varying Independent Variables	599
15.7.3	Extension to Multiple Episode Data	599
15.7.4	Extension to a Categorical Outcome: Event-History Analysis	600
15.8	Time Series Analysis	600
15.8.1	Units of Observation in Time Series Analyses	601
15.8.2	Time Series Analyses Applications	601
15.8.3	Time Effects in Time Series	602
15.8.4	Extension of Time Series Analyses to Multiple Units or Subjects	602
15.9	Dynamic System Analysis	602
15.10	Statistical Inference and Power Analysis in Longitudinal Analyses	604
15.11	Summary	605

Chapter 16: Multiple Dependent Variables: Set Correlation

608

16.1	Introduction to Ordinary Least Squares Treatment of Multiple Dependent Variables	608
16.1.1	Set Correlation Analysis	608

16.1.2	Canonical Analysis	609
16.1.3	Elements of Set Correlation	610
16.2	Measures of Multivariate Association	610
16.2.1	$R^2_{Y,X}$, the Proportion of Generalized Variance	610
16.2.2	$T^2_{Y,X}$ and $P^2_{Y,X}$, Proportions of Additive Variance	611
16.3	Partialing in Set Correlation	613
16.3.1	Frequent Reasons for Partialing Variable Sets From the Basic Sets	613
16.3.2	The Five Types of Association Between Basic Y and X Sets	614
16.4	Tests of Statistical Significance and Statistical Power	615
16.4.1	Testing the Null Hypothesis	615
16.4.2	Estimators of the Population $R^2_{Y,X}$, $T^2_{Y,X}$, and $P^2_{Y,X}$	616
16.4.3	Guarding Against Type I Error Inflation	617
16.5	Statistical Power Analysis in Set Correlation	617
16.6	Comparison of Set Correlation With Multiple Analysis of Variance	619
16.7	New Analytic Possibilities With Set Correlation	620
16.8	Illustrative Examples	621
16.8.1	A Simple Whole Association	621
16.8.2	A Multivariate Analysis of Partial Variance	622
16.8.3	A Hierarchical Analysis of a Quantitative Set and Its Unique Components	623
16.8.4	Bipartial Association Among Three Sets	625
16.9	Summary	627

APPENDICES

Appendix 1: The Mathematical Basis for Multiple Regression/Correlation and Identification of the Inverse Matrix Elements 631

A1.1	Alternative Matrix Methods	634
A1.2	Determinants	634

Appendix 2: Determination of the Inverse Matrix and Applications Thereof 636

A2.1	Hand Calculation of the Multiple Regression/Correlation Problem	636
A2.2	Testing the Difference Between Partial β s and B s From the Same Sample	640
A2.3	Testing the Difference Between β s for Different Dependent Variables From a Single Sample	642

Appendix Tables **643**

Table A *t* Values for $\alpha = .01, .05$ (Two Tailed) 643

Table B z' Transformation of r 644

Table C Normal Distribution 645

Table D F Values for $\alpha = .01, .05$ 646

Table E L Values for $\alpha = .01, .05$ 650

Table F Power of Significance Test of r at $\alpha = .01, .05$
 (Two Tailed) 652

Table G n^* to Detect r by t Test at $\alpha = .01, .05$
 (Two Tailed) 654

References **655**

Glossary **671**

Statistical Symbols and Abbreviations **683**

Author Index **687**

Subject Index **691**