Contents

Preface xiv
Acknowledgments xvii

1 Overview of Multivariate Methods 1

What Is Multivariate Analysis? 3

Three Converging Trends 4
Topic 1: Rise of Big Data 4
Topic 2: Statistical Versus Data Mining Models 7
Topic 3: Causal Inference 9
Summary 9

Multivariate Analysis in Statistical Terms 9

Some Basic Concepts of Multivariate Analysis 10
The Variate 10
Measurement Scales 11
Measurement Error and Multivariate Measurement 13

Managing the Multivariate Model 14
Managing the Variate 14
Managing the Dependence Model 17
Statistical Significance Versus Statistical Power 18
Review 20

A Classification of Multivariate Techniques 21
Dependence Techniques 21
Interdependence Techniques 25

Types of Multivariate Techniques 25
Exploratory Factor Analysis: Principal Components and Common Factor Analysis 25
Cluster Analysis 26
Multiple Regression 26
Multivariate Analysis of Variance and Covariance 26
Multiple Discriminant Analysis 26
Logistic Regression 27
Structural Equation Modeling and Confirmatory Factor Analysis 27
Partial Least Squares Structural Equation Modeling 28
Canonical Correlation 28
Conjoint Analysis 28
Perceptual Mapping 29
Correspondence Analysis 29

Guidelines for Multivariate Analyses and Interpretation 29
Establish Practical Significance as Well as Statistical Significance 30
Recognize That Sample Size Affects All Results 30
Know Your Data 30
Strive for Model Parsimony 31
Look at Your Errors 31
Simplify Your Models By Separation 31
Validate Your Results 32

A Structured Approach to Multivariate Model Building 32
Stage 1: Define the Research Problem, Objectives, and Multivariate Technique to Be Used 33
Stage 2: Develop the Analysis Plan 33
Stage 3: Evaluate the Assumptions Underlying the Multivariate Technique 33
Stage 4: Estimate the Multivariate Model and Assess Overall Model Fit 34
Stage 5: Interpret the Variate(s) 34
Stage 6: Validate the Multivariate Model 34
A Decision Flowchart 34

Databases 34
Primary Database 35
Other Databases 37

Organization of the Remaining Chapters 37
Section I: Preparing for a Multivariate Analysis 37
Section II: Interdependence Techniques 38
Sections III and IV: Dependence Techniques 38
Section V: Moving Beyond the Basics 38
Online Resources: Additional Chapters 38

Summary 39

Questions 41

Suggested Readings and Online Resources 41

References 41

SECTION I
Preparing for Multivariate Analysis 43

2 Examining Your Data 45

Introduction 49

The Challenge of Big Data Research Efforts 49

Data Management 50
Data Quality 50
Summary 51

Preliminary Examination of the Data 51
Univariate Profiling: Examining the Shape of the Distribution 51
Bivariate Profiling: Examining the Relationship Between Variables 52
Bivariate Profiling: Examining Group Differences 53
Multivariate Profiles 54
New Measures of Association 55
Summary 55

Missing Data 56
The Impact of Missing Data 56
Recent Developments in Missing Data Analysis 57
A Simple Example of a Missing Data Analysis 57
A Four-Step Process for Identifying Missing Data and Applying Remedies 58
An Illustration of Missing Data Diagnosis with the Four-Step Process 72

Outliers 85
Two Different Contexts for Defining Outliers 85
Impacts of Outliers 86
Classifying Outliers 87
Detecting and Handling Outliers 88
An Illustrative Example of Analyzing Outliers 91

Testing the Assumptions of Multivariate Analysis 93
Assessing Individual Variables Versus the Variate 93
Four Important Statistical Assumptions 94

Data Transformations 100
Transformations Related to Statistical Properties 101
Transformations Related to Interpretation 101
Transformations Related to Specific Relationship Types 102
Transformations Related to Simplification 103
General Guidelines for Transformations 104

An Illustration of Testing the Assumptions Underlying Multivariate Analysis 105
Normality 105
Homoscedasticity 108
Linearity 108
Summary 112

Incorporating Nonmetric Data with Dummy Variables 112
Concept of Dummy Variables 112
Dummy Variable Coding 113
Using Dummy Variables 113

Summary 114

Questions 115

Suggested Readings and Online Resources 116

References 116

SECTION II
Interdependence Techniques 119

3 Exploratory Factor Analysis 121

What Is Exploratory Factor Analysis? 124

A Hypothetical Example of Exploratory Factor Analysis 126

Factor Analysis Decision Process 127

Stage 1: Objectives of Factor Analysis 127
Specifying the Unit of Analysis 127
Achieving Data Summarization Versus Data Reduction 129
Variable Selection 131
Using Factor Analysis with Other Multivariate Techniques 131

Stage 2: Designing an Exploratory Factor Analysis 132
Variable Selection and Measurement Issues 132
Sample Size 132
Correlations among Variables or Respondents 133

Stage 3: Assumptions in Exploratory Factor Analysis 135
Conceptual Issues 135
Statistical Issues 135
Summary 136

Stage 4: Deriving Factors and Assessing Overall Fit 136
Selecting the Factor Extraction Method 138
Stopping Rules: Criteria for the Number of Factors to Extract 140
Alternatives to Principal Components and Common Factor Analysis 144

Stage 5: Interpreting the Factors 146
The Three Processes of Factor Interpretation 146
Factor Extraction 147
Rotation of Factors 147
Judging the Significance of Factor Loadings 151
Interpreting a Factor Matrix 153

Stage 6: Validation of Exploratory Factor Analysis 158
Use of Replication or a Confirmatory Perspective 158
Assessing Factor Structure Stability 159
Detecting Influential Observations 159

Stage 7: Data Reduction—Additional Uses of Exploratory Factor Analysis Results 159
Selecting Surrogate Variables for Subsequent Analysis 160
Creating Summated Scales 160
Computing Factor Scores 163
Selecting among the Three Methods 164
An Illustrative Example 165
Stage 1: Objectives of Factor Analysis 165
Stage 2: Designing a Factor Analysis 165
Stage 3: Assumptions in Factor Analysis 165
Principal Component Factor Analysis: Stages 4–7 168
Common Factor Analysis: Stages 4 and 5 181
A Managerial Overview of the Results 183
Summary 184
Questions 187
Suggested Readings and Online Resources 187
References 187

4 Cluster Analysis 189

What Is Cluster Analysis? 192
Cluster Analysis as a Multivariate Technique 192
Conceptual Development with Cluster Analysis 192
Necessity of Conceptual Support in Cluster Analysis 193
How Does Cluster Analysis Work? 193
A Simple Example 194
Objective Versus Subjective Considerations 199
Cluster Analysis Decision Process 199
Stage 1: Objectives of Cluster Analysis 199
Stage 2: Research Design in Cluster Analysis 202
Stage 3: Assumptions in Cluster Analysis 211
Stage 4: Deriving Clusters and Assessing Overall Fit 212
Stage 5: Interpretation of the Clusters 227
Stage 6: Validation and Profiling of the Clusters 228
Implication of Big Data Analytics 230
Challenges 230
An Illustrative Example 230
Stage 1: Objectives of the Cluster Analysis 231
Stage 2: Research Design of the Cluster Analysis 232
Stage 3: Assumptions in Cluster Analysis 235
Stages 4–6: Employing Hierarchical and Nonhierarchical Methods 235
Part 1: Hierarchical Cluster Analysis (Stage 4) 235
Part 2: Nonhierarchical Cluster Analysis (Stages 4–6) 245
Examining an Alternative Cluster Solution: Stages 4–6 251
A Managerial Overview of the Clustering Process 252
Summary 253
Questions 254
Suggested Readings and Online Resources 255
References 255

SECTION III
Dependence Techniques – Metric Outcomes 257

5 Multiple Regression Analysis 259

What Is Multiple Regression Analysis? 265
Multiple Regression in the Era of Big Data 265
An Example of Simple and Multiple Regression 266
Prediction Using a Single Independent Variable: Simple Regression 267
Prediction Using Several Independent Variables: Multiple Regression 269
Summary 271
A Decision Process for Multiple Regression Analysis 272
Stage 1: Objectives of Multiple Regression 273
Research Problems Appropriate for Multiple Regression 273
Specifying a Statistical Relationship 274
Selection of Dependent and Independent Variables 275
Stage 2: Research Design of a Multiple Regression Analysis 278
Sample Size 278
Creating Additional Variables 281
Overview 286
Stage 3: Assumptions in Multiple Regression Analysis 287
Assessing Individual Variables Versus the Variate 287
Methods of Diagnosis 288
Linearity of the Phenomenon 288
Constant Variance of the Error Term 290
Normality of the Error Term Distribution 291
Independence of the Error Terms 291
Summary 292
Stage 4: Estimating the Regression Model and Assessing Overall Model Fit 292
Managing the Variate 292

Variable Specification 294
Variable Selection 295
Testing the Regression Variate for Meeting the Regression Assumptions 298
Examining the Statistical Significance of Our Model 299
Understanding Influential Observations 302

Stage 5: Interpreting the Regression Variate 308
Using the Regression Coefficients 308
Assessing Multicollinearity 311
Relative Importance of Independent Variables 317
Summary 320

Stage 6: Validation of the Results 321
Additional or Split Samples 321
Calculating the PRESS Statistic 321
Comparing Regression Models 322
Forecasting with the Model 322

Extending Multiple Regression 322
Multilevel Models 323
Panel Models 328

Illustration of a Regression Analysis 331
Stage 1: Objectives of Multiple Regression 331
Stage 2: Research Design of a Multiple Regression Analysis 331
Stage 3: Assumptions in Multiple Regression Analysis 332
Stage 4: Estimating the Regression Model and Assessing Overall Model Fit 332
Stage 5: Interpreting the Regression Variate 348
Stage 6: Validating the Results 353

Evaluating Alternative Regression Models 355
Confirmatory Regression Model 355
Use of Summated Scales as Remedies for Multicollinearity 357
Including a Nonmetric Independent Variable 361
A Managerial Overview of the Results 361

Summary 363

Questions 366

Suggested Readings and Online Resources 367

References 367

6 MANOVA: Extending ANOVA 371

Re-Emergence of Experimentation 376

Experimental Approaches Versus Other Multivariate Methods 376

MANOVA: Extending Univariate Methods for Assessing Group Differences 377
Multivariate Procedures for Assessing Group Differences 377

A Hypothetical Illustration of MANOVA 381
Analysis Design 381
Differences from Discriminant Analysis 381
Forming the Variate and Assessing Differences 382

A Decision Process for MANOVA 383

Stage 1: Objectives of MANOVA 385
When Should We Use MANOVA? 385
Types of Multivariate Questions Suitable for MANOVA 385
Selecting the Dependent Measures 386

Stage 2: Issues in the Research Design of MANOVA 387
Types of Research Approaches 387
Types of Variables in Experimental Research 389
Sample Size Requirements—Overall and by Group 391
Factorial Designs—Two or More Treatments 391
Using Covariates—ANCOVA and MANCOVA 394
Modeling Other Relationships Between Treatment and Outcome 396
MANOVA Counterparts of Other ANOVA Designs 397
A Special Case of MANOVA: Repeated Measures 397

Stage 3: Assumptions of ANOVA and MANOVA 398
Independence 399
Equality of Variance–Covariance Matrices 399
Normality 400
Linearity and Multicollinearity among the Dependent Variables 401
Sensitivity to Outliers 401

Stage 4: Estimation of the MANOVA Model and Assessing Overall Fit 401
Estimation with the General Linear Model 403
Measures for Significance Testing 403
Statistical Power of the Multivariate Tests 403
Estimating Additional Relationships: Mediation and Moderation 407

Stage 5: Interpretation of the MANOVA Results 410
Evaluating Covariates 410
Assessing Effects on the Dependent Variate 411
Identifying Differences Between Individual Groups 415
Assessing Significance for Individual Outcome Variables 417
Interpreting Mediation and Moderation 419

Stage 6: Validation of the Results 421

Advanced Issues: Causal Inference in Nonrandomized Situations 421

Causality in the Social and Behavioral Sciences 422
The Potential Outcomes Approach 423
Counterfactuals in Non-experimental Research Designs 423
Propensity Score Models 424
Overview 428

Summary 430

Illustration of a MANOVA Analysis 430
Research Setting 430

Example 1: Difference Between Two Independent Groups 432
Stage 1: Objectives of the Analysis 432
Stage 2: Research Design of the MANOVA 433
Stage 3: Assumptions in MANOVA 433
Stage 4: Estimation of the MANOVA Model and Assessing Overall Fit 434
Stage 5: Interpretation of the Results 437
Summary 438

Example 2: Difference Between K Independent Groups 438
Stage 1: Objectives of the MANOVA 438
Stage 2: Research Design of MANOVA 439
Stage 3: Assumptions IN MANOVA 439
Stage 4: Estimation of the MANOVA Model and Assessing Overall Fit 440
Stage 5: Interpretation of the Results 443
Summary 444

Example 3: A Factorial Design for MANOVA with Two Independent Variables 444
Stage 1: Objectives of the MANOVA 445
Stage 2: Research Design of the MANOVA 445
Stage 3: Assumptions in MANOVA 447
Stage 4: Estimation of the MANOVA Model and Assessing Overall Fit 448
Stage 5: Interpretation of the Results 451
Summary 452

Example 4: Moderation and Mediation 452
Moderation of Distribution System (X_5) by Firm Size (X_3) 453
Summary 456
Mediation of Distribution System (X_5) By Purchase Level (X_{22}) 457
Summary 459

A Managerial Overview of the Results 459

Summary 460

Questions 463

Suggested Readings and Online Resources 464

References 464

SECTION IV
Dependence Techniques – Non-metric Outcomes 469

7 Multiple Discriminant Analysis 471

What Is Discriminant Analysis? 474
The Variate 474
Testing Hypotheses 475

Similarities to Other Multivariate Techniques 476

Hypothetical Example of Discriminant Analysis 476
A Two-Group Discriminant Analysis: Purchasers Versus Non-purchasers 476
A Three-Group Example of Discriminant Analysis: Switching Intentions 481

The Decision Process for Discriminant Analysis 484

Stage 1: Objectives of Discriminant Analysis 484
Descriptive Profile Analysis 485
Classification Purposes 485

Stage 2: Research Design for Discriminant Analysis 485
Selecting Dependent and Independent Variables 485
Sample Size 487
Division of the Sample 488

Stage 3: Assumptions of Discriminant Analysis 488
Impacts on Estimation and Classification 489
Impacts on Interpretation 489

Stage 4: Estimation of the Discriminant Model and Assessing Overall Fit 490
Selecting an Estimation Method 491
Statistical Significance 492
Assessing Overall Model Fit 493
Casewise Diagnostics 501

Stage 5: Interpretation of the Results 503
Discriminant Weights 503
Discriminant Loadings 503
Partial *F* Values 504
Interpretation of Two or More Functions 504
Which Interpretive Method to Use? 506

Stage 6: Validation of the Results 506
Validation Procedures 506
Profiling Group Differences 507

A Two-Group Illustrative Example 508
Stage 1: Objectives of Discriminant Analysis 508
Stage 2: Research Design for Discriminant Analysis 508
Stage 3: Assumptions of Discriminant Analysis 509
Stage 4: Estimation of the Discriminant Model and Assessing Overall Fit 509

Stage 5: Interpretation of the Results 520
Stage 6: Validation of the Results 522
A Managerial Overview 523

A Three-Group Illustrative Example 523
Stage 1: Objectives of Discriminant Analysis 524
Stage 2: Research Design for Discriminant Analysis 524
Stage 3: Assumptions of Discriminant Analysis 524
Stage 4: Estimation of the Discriminant Model and Assessing Overall Fit 525
Stage 5: Interpretation of Three-Group Discriminant Analysis Results 537
Stage 6: Validation of the Discriminant Results 542
A Managerial Overview 543

Summary 544

Questions 546

Suggested Readings and Online Resources 547

References 547

8 Logistic Regression: Regression with a Binary Dependent Variable 548

What Is Logistic Regression? 551

The Decision Process for Logistic Regression 552

Stage 1: Objectives of Logistic Regression 552
Explanation 552
Classification 553

Stage 2: Research Design for Logistic Regression 553
Representation of the Binary Dependent Variable 553
Sample Size 555
Use of Aggregated Data 556

Stage 3: Assumptions of Logistic Regression 556

Stage 4: Estimation of the Logistic Regression Model and Assessing Overall Fit 557
Estimating the Logistic Regression Model 558
Assessing the Goodness-of-Fit of the Estimated Model 563
Overview of Assessing Model Fit 571
Casewise Diagnostics 571
Summary 572

Stage 5: Interpretation of the Results 572
Testing for Significance of the Coefficients 573
Interpreting the Coefficients 574
Calculating Probabilities for a Specific Value of the Independent Variable 578
Overview of Interpreting Coefficients 579

Stage 6: Validation of the Results 579

An Illustrative Example of Logistic Regression 580
Stage 1: Objectives of Logistic Regression 580
Stage 2: Research Design for Logistic Regression 580
Stage 3: Assumptions of Logistic Regression 581
Stage 4: Estimation of the Logistic Regression Model and Assessing Overall Fit 581
Stage 5: Interpretation of Results 592
Stage 6: Validation of the Results 596
A Managerial Overview 596

Summary 596

Questions 598

Suggested Readings and Online Resources 598

References 598

SECTION V
Moving Beyond The Basics 601

9 Structural Equation Modeling: An Introduction 603

What Is Structural Equation Modeling? 607
Estimation of Multiple Interrelated Dependence Relationships 607
Incorporating Latent Variables Not Measured Directly 608
Defining a Model 610

SEM and Other Multivariate Techniques 613
Similarity to Dependence Techniques 613
Similarity to Interdependence Techniques 613
The Emergence of SEM 614

The Role of Theory in Structural Equation Modeling 614
Specifying Relationships 614
Establishing Causation 615
Developing a Modeling Strategy 618

A Simple Example of SEM 619
Theory 619
Setting Up the Structural Equation Model for Path Analysis 620
The Basics of SEM Estimation and Assessment 621

Six Stages in Structural Equation Modeling 625

Stage 1: Defining Individual Constructs 627
Operationalizing the Construct 627
Pretesting 627

Stage 2: Developing and Specifying the Measurement Model 627
SEM Notation 628
Creating the Measurement Model 629
Stage 3: Designing a Study to Produce Empirical Results 629
Issues in Research Design 629
Issues in Model Estimation 633
Stage 4: Assessing Measurement Model Validity 635
The Basics of Goodness-of-Fit 635
Absolute Fit Indices 636
Incremental Fit Indices 638
Parsimony Fit Indices 639
Problems Associated with Using Fit Indices 639
Unacceptable Model Specification to Achieve Fit 641
Guidelines for Establishing Acceptable and Unacceptable Fit 641
Stage 5: Specifying the Structural Model 643
Stage 6: Assessing the Structural Model Validity 644
Competitive Fit 645
Testing Structural Relationships 647
Summary 648
Questions 649
Suggested Readings and Online Resources 649
Appendix 9A: Estimating Relationships Using Path Analysis 650
Appendix 9B: SEM Abbreviations 653
Appendix 9C: Detail on Selected GOF Indices 654
References 656

10 SEM: Confirmatory Factor Analysis 658

What Is Confirmatory Factor Analysis? 660
CFA and Exploratory Factor Analysis 660
Measurement Theory and Psychometrics 661
A Simple Example of CFA and SEM 661
A Visual Diagram 661
SEM Stages for Testing Measurement Theory Validation with CFA 663
Stage 1: Defining Individual Constructs 663
Stage 2: Developing the Overall Measurement Model 663
Unidimensionality 664
Congeneric Measurement Model 665
Items per Construct 665
Reflective Versus Formative Measurement 668
Stage 3: Designing a Study to Produce Empirical Results 670
Measurement Scales in CFA 670
SEM and Sampling 670
Specifying the Model 670
Issues in Identification 671
Problems in Estimation 673
Stage 4: Assessing Measurement Model Validity 673
Assessing Fit 674
Path Estimates 674
Construct Validity 675
Model Diagnostics 677
Summary Example 681
CFA Illustration 681
Stage 1: Defining Individual Constructs 682
Stage 2: Developing the Overall Measurement Model 682
Stage 3: Designing a Study to Produce Empirical Results 684
Stage 4: Assessing Measurement Model Validity 685
HBAT CFA Summary 692
CFA Results Detect Problems 693
Summary 696
Questions 697
Suggested Readings and Online Resources 697
References 697

11 Testing Structural Equation Models 699

What Is a Structural Model? 700
A Simple Example of a Structural Model 701
An Overview of Theory Testing with SEM 702
Stages in Testing Structural Theory 703
One-Step Versus Two-Step Approaches 703
Stage 5: Specifying the Structural Model 703
Unit of Analysis 704
Model Specification Using a Path Diagram 704
Designing the Study 708
Stage 6: Assessing the Structural Model Validity 710
Understanding Structural Model Fit from CFA Fit 710
Examine the Model Diagnostics 712
SEM Illustration 713
Stage 5: Specifying the Structural Model 713
Stage 6: Assessing the Structural Model Validity 715
Summary 722
Questions 723
Suggested Readings and Online Resources 723

Appendix 11A 724
References 725

12 Advanced SEM Topics 726

Reflective Versus Formative Scales 728
Reflective Versus Formative Measurement Theory 728
Operationalizing a Formative Measure 729
Differences Between Reflective and Formative Measures 730
Which to Use—Reflective or Formative? 732

Higher-Order Factor Models 732
Empirical Concerns 733
Theoretical Concerns 734
Using Second-Order Measurement Theories 735
When to Use Higher-Order Factor Analysis 736

Multiple Groups Analysis 736
Measurement Model Comparisons 737
Structural Model Comparisons 741

Measurement Type Bias 742
Model Specification 742
Model Interpretation 744

Relationship Types: Mediation and Moderation 744
Mediation 745
Moderation 748

Developments in Advanced SEM Approaches 752
Longitudinal Data 752
Latent Growth Models 752
Bayesian SEM 753

Summary 755

Questions 756

Suggested Readings and Online Resources 757

References 757

13 Partial Least Squares Structural Equation Modeling (PLS-SEM) 759

What is PLS-SEM? 764
Structural Model 764
Measurement Model 764
Theory and Path Models in PLS-SEM 765
The Emergence of SEM 765
Role of PLS-SEM Versus CB-SEM 766

Estimation of Path Models with PLS-SEM 766
Measurement Model Estimation 766
Structural Model Estimation 767
Estimating the Path Model Using the PLS-SEM Algorithm 767

PLS-SEM Decision Process 768

Stage 1: Defining Research Objectives and Selecting Constructs 768

Stage 2: Designing a Study to Produce Empirical Results 769
Metric Versus Nonmetric Data and Multivariate Normality 769
Missing Data 770
Statistical Power 770
Model Complexity and Sample Size 770

Stage 3: Specifying the Measurement and Structural Models 771
Measurement Theory and Models 773
Structural Theory and Path Models 774

Stage 4: Assessing Measurement Model Validity 774
Assessing Reflective Measurement Models 775
Assessing Formative Measurement Models 776
Summary 779

Stage 5: Assessing the Structural Model 779
Collinearity among Predictor Constructs 779
Examining the Coefficient of Determination 780
Effect Size 780
Blindfolding 780
Size and Significance of Path Coefficients 780
Summary 781

Stage 6: Advanced Analyses Using PLS-SEM 782
Multi-Group Analysis of Observed Heterogeneity 782
Detecting Unobserved Heterogeneity 782
Confirmatory Tetrad Analysis 782
Mediation Effects 782
Moderation 783
Higher-Order Measurement Models 783
Summary 783

PLS-SEM Illustration 783
Theoretical PLS-SEM Path Model 784

Stage 4: Assessing Measurement Model Reliability and Validity 785
Path Coefficients 785
Construct Reliability 786
Construct Validity 787
HBAT CCA Summary 790

Stage 5: Assessing the Structural Model 790
HBAT PLS-SEM Summary 791

Summary 792

Questions 793

Suggested Readings and Online Resources 793

References 793

Index 800