

Contents

Preface xv

Chapter 1	Introduction	1
1.1	MULTIVARIATE STATISTICS: WHY?	1
1.1.1	The Domain of Multivariate Statistics: Numbers of IVs and DVs	2
1.1.2	Experimental and Nonexperimental Research	3
1.1.2.1	Multivariate Statistics in Nonexperimental Research	4
1.1.2.2	Multivariate Statistics in Experimental Research	5
1.1.3	Computers and Multivariate Statistics	5
1.1.3.1	Program Updates	7
1.1.3.2	Garbage In, Roses Out?	7
1.1.4	Why Not?	8
1.2	SOME USEFUL DEFINITIONS	8
1.2.1	Continuous, Discrete, and Dichotomous Data	9
1.2.2	Samples and Populations	10
1.2.3	Descriptive and Inferential Statistics	11
1.2.4	Orthogonality	11
1.2.5	Standard and Hierarchical Analyses	12
1.3	KINDS OF RESEARCH QUESTIONS	13
1.3.1	Measuring Degree of Relationship Among Variables	13
1.3.2	Measuring the Significance of Group Differences	14
1.3.3	Predicting Group Membership	15
1.3.4	Determining Structure	15
1.4	COMBINING VARIABLES	15
1.5	NUMBER AND NATURE OF VARIABLES TO INCLUDE	17

1.6 DATA APPROPRIATE FOR MULTIVARIATE STATISTICS	17
1.6.1 The Data Matrix	18
1.6.2 The Correlation Matrix	18
1.6.3 The Variance-Covariance Matrix	19
1.6.4 The Sum-of-Squares and Cross-Products Matrix	20
1.6.5 Residuals	22
1.7 ORGANIZATION OF THE BOOK	22
Chapter 2 Review of Univariate and Bivariate Statistics	24
2.1 HYPOTHESIS TESTING	24
2.1.1 One-Sample z Test as Prototype	24
2.1.2 Power	28
2.1.3 Extensions of the Model	28
2.2 ANALYSIS OF VARIANCE	29
2.2.1 One-Way Between-Subjects ANOVA	30
2.2.2 Factorial Between-Subjects ANOVA	34
2.2.3 Within-Subjects ANOVA	35
2.2.4 Mixed Between-Within-Subjects ANOVA	38
2.2.5 Design Complexity	40
2.2.5.1 Nesting	40
2.2.5.2 Latin Square Designs	41
2.2.5.3 Unequal n and Nonorthogonality	41
2.2.5.4 Fixed and Random Effects	42
2.2.6 Specific Comparisons	42
2.2.6.1 Post Hoc Comparisons	42
2.2.6.2 Planned Comparisons	43
2.2.6.3 Weighting Coefficients	43
2.2.6.4 Orthogonality	44
2.2.6.5 F Tests for Comparisons	44
2.2.6.6 Type IV Errors	45
2.3 PARAMETER ESTIMATION	46
2.4 STRENGTH OF ASSOCIATION	46
2.5 BIVARIATE STATISTICS: CORRELATION AND REGRESSION	48
2.5.1 Correlation	48
2.5.2 Regression	49
Chapter 3 A Guide to Statistical Techniques: Using the Book	52
3.1 RESEARCH QUESTIONS AND ASSOCIATED TECHNIQUES	52
3.1.1 Degree of Relationship Among Variables	52
3.1.1.1 Bivariate r	53
3.1.1.2 Multiple R	53
3.1.1.3 Hierarchical R	53
3.1.1.4 Canonical R	54
3.1.2 Significance of Group Differences	54
3.1.2.1 One-Way ANOVA and t Test	55
3.1.2.2 One-Way ANCOVA	55

3.1.2.3	Factorial ANOVA	56	
3.1.2.4	Factorial ANCOVA	56	
3.1.2.5	Hotelling's T^2	56	
3.1.2.6	One-Way MANOVA	57	
3.1.2.7	One-Way MANCOVA	58	
3.1.2.8	Factorial MANOVA	58	
3.1.2.9	Factorial MANCOVA	59	
3.1.3	Prediction of Group Membership	59	
3.1.3.1	One-Way Discriminant Function	59	
3.1.3.2	Hierarchical One-Way Discriminant Function		60
3.1.3.3	Factorial Discriminant Function	60	
3.1.3.4	Hierarchical Factorial Discriminant Function		60
3.1.4	Structure	60	
3.1.4.1	Principal Components	60	
3.1.4.2	Factor Analysis	62	
3.2	A DECISION TREE		62
3.3	TECHNIQUE CHAPTERS		64
3.4	PRELIMINARY CHECK OF THE DATA		65
Chapter 4	Conditioning Matrices: Cleaning up Your Act Before Analyzing Your Data		66
4.1	HONEST CORRELATIONS		67
4.1.1	Inflated Correlation	67	
4.1.2	Deflated Correlation	67	
4.1.3	Inaccurately Computed R	68	
4.2	MISSING DATA		68
4.2.1	Treating Missing Data as Data	69	
4.2.2	Deleting Cases or Variables	69	
4.2.3	Using a Missing Data Correlation Matrix	70	
4.2.4	Estimating Missing Data	71	
4.2.4.1	Using Prior Knowledge	71	
4.2.4.2	Inserting Mean Values	71	
4.2.4.3	Using Regression to Predict Missing Values	72	
4.2.5	Repeating Analyses With and Without Missing Data	72	
4.3	OUTLIERS		72
4.3.1	Detecting Outliers	74	
4.3.1.1	Univariate Outliers	74	
4.3.1.2	Multivariate Outliers	74	
4.3.2	Reducing the Influence of Outliers	76	
4.3.2.1	Checking Accuracy of Data	76	
4.3.2.2	Transforming Distributions with Outliers	76	
4.3.2.3	Deleting Outlying Cases	76	
4.3.2.4	Deleting Variables	76	
4.3.2.5	Changing Scores on Variables Causing Outliers	76	
4.4	ASSUMPTIONS		76
4.4.1	Normality	77	
4.4.2	Linearity	80	
4.4.3	Homoscedasticity	81	
4.4.4	Multicollinearity and Singularity	82	

4.5 COMMON DATA TRANSFORMATIONS	84
4.6 PRELIMINARY EVALUATION OF DATA	85
Chapter 5 Multiple Regression	86
5.1 GENERAL PURPOSE AND DESCRIPTION	86
5.2 KINDS OF RESEARCH QUESTIONS	87
5.2.1 Degree of Relationship	88
5.2.2 Importance of IVs	88
5.2.3 Adding IVs	89
5.2.4 Changing IVs	89
5.2.5 Contingencies Among IVs	89
5.2.6 Comparing Sets of IVs	89
5.2.7 Predicting DV Scores for Members of a New Sample	90
5.2.8 Causal Modeling	90
5.3 LIMITATIONS TO REGRESSION ANALYSES	91
5.3.1 Theoretical Issues	91
5.3.2 Practical Issues	91
5.3.2.1 Number of Cases and Variables	91
5.3.2.2 Outliers	92
5.3.2.3 Multicollinearity and Singularity	92
5.3.2.4 Normality, Linearity, and Homoscedasticity of Residuals	93
5.4 FUNDAMENTAL EQUATIONS FOR MULTIPLE REGRESSION	95
5.4.1 General Linear Equation	96
5.4.2 Matrix Equations	98
5.5 MAJOR TYPES OF MULTIPLE REGRESSION	101
5.5.1 Standard Multiple Regression	101
5.5.2 Hierarchical Multiple Regression	102
5.5.3 Stepwise and Setwise Regression	103
5.5.4 Choosing Among Regression Strategies	105
5.6 SOME IMPORTANT ISSUES	107
5.6.1 Importance of IVs	107
5.6.1.1 Standard Multiple and Setwise Regression	108
5.6.1.2 Hierarchical or Stepwise Regression	109
5.6.2 Statistical Inference	109
5.6.2.1 Test of Multiple R	109
5.6.2.2 Test of Regression Components	111
5.6.2.3 Test of Added Subset of IVs	112
5.6.2.4 Confidence Limits Around B	113
5.6.2.5 Comparing Two Sets of Predictors	114
5.6.3 Adjustment of R^2	115
5.6.4 Suppressor Variables	116
5.7 COMPARISON OF SPSS AND BMD PROGRAMS	117
5.7.1 SPSS Package	117
5.7.2 BMD Series	125
5.8 COMPLETE EXAMPLES OF REGRESSION ANALYSIS	126
5.8.1 Evaluation of Assumptions	126
5.8.1.1 Number of Cases and Variables	126
5.8.1.2 Normality, Linearity, and Homoscedasticity of Residuals	126

5.8.1.3 Outliers	129	
5.8.1.4 Multicollinearity and Singularity	132	
5.8.2 Standard Multiple Regression	132	
5.8.3 Hierarchical Regression	139	
5.9 SOME EXAMPLES FROM THE LITERATURE		144
Chapter 6 Canonical Correlation		146
6.1 GENERAL PURPOSE AND DESCRIPTION		146
6.2 KINDS OF RESEARCH QUESTIONS		147
6.2.1 Number of Canonical Variate Pairs	147	
6.2.2 Nature of Canonical Variates	147	
6.2.3 Importance of Canonical Variates	148	
6.2.4 Canonical Variate Scores	148	
6.3 LIMITATIONS		148
6.3.1 Theoretical Limitations	148	
6.3.2 Practical Issues	149	
6.3.2.1 Normality	149	
6.3.2.2 Missing Data	149	
6.3.2.3 Outliers	149	
6.3.2.4 Multicollinearity and Singularity	150	
6.3.2.5 Linearity	150	
6.4 FUNDAMENTAL EQUATIONS FOR CANONICAL CORRELATION		150
6.5 SOME IMPORTANT ISSUES		157
6.5.1 Importance of Canonical Variates	157	
6.5.2 Interpretation of Canonical Variates	159	
6.6 COMPARISON OF SPSS AND BMDP PROGRAMS		159
6.6.1 SPSS Package	159	
6.6.2 BMDP Series	161	
6.7 COMPLETE EXAMPLE OF CANONICAL CORRELATION		161
6.7.1 Evaluation of Assumptions	162	
6.7.1.1 Normality	162	
6.7.1.2 Missing Data	162	
6.7.1.3 Outliers	162	
6.7.1.4 Multicollinearity and Singularity	164	
6.7.1.5 Linearity	164	
6.7.2 Canonical Correlation	165	
6.8 SOME EXAMPLES FROM THE LITERATURE		171
Chapter 7 Analysis of Covariance		173
7.1 GENERAL PURPOSE AND DESCRIPTION		173
7.2 KINDS OF RESEARCH QUESTIONS		176
7.2.1 Main Effects of IVs	176	
7.2.2 Interactions Among IVs	177	
7.2.3 Specific Comparisons and Trend Analysis	177	
7.2.4 Effects of Covariates	178	
7.2.5 Strength of Association	178	
7.2.6 Adjusted Marginal and Cell Means	178	

7.3 LIMITATIONS TO ANALYSIS OF COVARIANCE	179
7.3.1 Theoretical Issues	179
7.3.2 Practical Issues	180
7.3.2.1 Unequal Sample Sizes and Missing Data	180
7.3.2.2 Outliers	180
7.3.2.3 Multicollinearity and Singularity	181
7.3.2.4 Normality	181
7.3.2.5 Homogeneity of Variance	181
7.3.2.6 Linearity	182
7.3.2.7 Homogeneity of Regression	182
7.3.2.8 Reliability of Covariates	182
7.4 FUNDAMENTAL EQUATIONS FOR ANALYSIS OF COVARIANCE	184
7.5 SOME IMPORTANT ISSUES	190
7.5.1 Test for Homogeneity of Regression	190
7.5.1.1 BMD Series	191
7.5.1.2 SPSS Package	192
7.5.2 Design Complexity	194
7.5.2.1 Within-Subjects and Mixed Within-Between Designs	194
7.5.2.2 Unequal Sample Sizes	195
7.5.2.3 Specific Comparisons and Trend Analysis	198
7.5.2.4 Strength of Association	201
7.5.3 Evaluation of Covariates	203
7.5.4 Choosing Covariates	204
7.5.5 Alternatives to ANCOVA	204
7.6 COMPARISON OF PROGRAMS	206
7.6.1 BMD Series	206
7.6.2 SPSS Package	209
7.7 COMPLETE EXAMPLE OF ANALYSIS OF COVARIANCE	209
7.7.1 Evaluation of Assumptions	210
7.7.1.1 Unequal n and Missing Data	210
7.7.1.2 Outliers	210
7.7.1.3 Normality	213
7.7.1.4 Multicollinearity and Singularity	213
7.7.1.5 Homogeneity of Variance	213
7.7.1.6 Linearity	213
7.7.1.7 Homogeneity of Regression	213
7.7.1.8 Reliability of Covariates	213
7.7.2 Analysis of Covariance	213
7.8 SOME EXAMPLES FROM THE LITERATURE	220
 Chapter 8 Multivariate Analysis of Variance and Covariance	 222
8.1 GENERAL PURPOSE AND DESCRIPTION	222
8.2 KINDS OF RESEARCH QUESTIONS	225
8.2.1 Main Effects of IVs	226
8.2.2 Interactions Among IVs	226
8.2.3 Importance of DVs	227

8.2.4	Specific Comparisons and Trend Analysis	227	
8.2.5	Strength of Association	227	
8.2.6	Repeated-Measures Analysis of Variance	228	
8.2.7	Adjusted Marginal and Cell Means	229	
8.2.8	Effects of Covariates	229	
8.3	LIMITATIONS TO MULTIVARIATE ANALYSIS OF VARIANCE AND COVARIANCE		230
8.3.1	Theoretical Issues	230	
8.3.2	Practical Issues	231	
8.3.2.1	Unequal Sample Sizes and Missing Data	231	
8.3.2.2	Multivariate Normality	232	
8.3.2.3	Outliers	232	
8.3.2.4	Homogeneity of Variance-Covariance Matrices	232	
8.3.2.5	Linearity	233	
8.3.2.6	Homogeneity of Regression	233	
8.3.2.7	Reliability of Covariates	234	
8.3.2.8	Multicollinearity and Singularity	234	
8.4	FUNDAMENTAL EQUATIONS FOR MULTIVARIATE ANALYSIS OF VARIANCE AND COVARIANCE		235
8.4.1	Multivariate Analysis of Variance	235	
8.4.2	Multivariate Analysis of Covariance	244	
8.5	SOME IMPORTANT ISSUES		248
8.5.1	Criteria for Statistical Inference	248	
8.5.2	Specific Comparisons and Trend Analysis	250	
8.5.3	Assessing DVs	252	
8.5.3.1	Univariate F	253	
8.5.3.2	Stepdown Analysis	254	
8.5.3.3	Standardized Discriminant Function Coefficients	256	
8.5.3.4	Loading Matrices in MANOVA	256	
8.5.3.5	Choosing Among Strategies for Assessing DVs	256	
8.5.4	Design Complexity	257	
8.5.4.1	Within-Subjects and Mixed Between-Within Designs	258	
8.5.4.2	Unequal Sample Sizes	259	
8.6	COMPARISON OF PROGRAMS		259
8.6.1	SPSS Package	259	
8.6.2	BMD Series	262	
8.7	COMPLETE EXAMPLES OF MULTIVARIATE ANALYSIS OF VARIANCE AND COVARIANCE		264
8.7.1	Evaluation of Assumptions	265	
8.7.1.1	Unequal Sample Sizes and Missing Data	265	
8.7.1.2	Outliers	265	
8.7.1.3	Multivariate Normality	265	
8.7.1.4	Homogeneity of Variance-Covariance Matrices	266	
8.7.1.5	Linearity	267	
8.7.1.6	Homogeneity of Regression	268	
8.7.1.7	Reliability of Covariates	269	
8.7.1.8	Multicollinearity and Singularity	270	
8.7.2	Multivariate Analysis of Variance	270	

8.7.3	Multivariate Analysis of Covariance	279
8.7.3.1	Assessing Covariates	279
8.7.3.2	Assessing DVs	282
8.8	SOME EXAMPLES FROM THE LITERATURE	288
8.8.1	Examples of MANOVA	288
8.8.2	Examples of MANCOVA	290
Chapter 9	Discriminant Function Analysis	292
9.1	GENERAL PURPOSE AND DESCRIPTION	292
9.2	KINDS OF RESEARCH QUESTIONS	294
9.2.1	Significance of Prediction	295
9.2.2	Number of Significant Discriminant Functions	295
9.2.3	Dimensions of Discrimination	296
9.2.4	Classification Functions	296
9.2.5	Adequacy of Classification	296
9.2.6	Strength of Association	297
9.2.7	Importance of Predictor Variables	297
9.2.8	Significance of Prediction with Covariates	298
9.2.9	Group Means	298
9.3	LIMITATIONS TO DISCRIMINANT FUNCTION ANALYSIS	298
9.3.1	Theoretical Issues	298
9.3.2	Practical Issues	299
9.3.2.1	Unequal Sample Sizes and Missing Data	299
9.3.2.2	Multivariate Normality	300
9.3.2.3	Outliers	300
9.3.2.4	Homogeneity of Variance-Covariance Matrices	300
9.3.2.5	Linearity	301
9.3.2.6	Multicollinearity and Singularity	301
9.4	FUNDAMENTAL EQUATIONS FOR DISCRIMINANT FUNCTION ANALYSIS	301
9.4.1	Test and Derivation of Discriminant Functions	302
9.4.2	Classification	306
9.5	TYPES OF DISCRIMINANT FUNCTION ANALYSIS	308
9.5.1	Direct Discriminant Function Analysis	309
9.5.2	Hierarchical Discriminant Function Analysis	309
9.5.3	Stepwise Discriminant Function Analysis	313
9.6	SOME IMPORTANT ISSUES	316
9.6.1	Statistical Inference	316
9.6.1.1	Criteria for Statistical Inference	316
9.6.1.2	Stepping Methods	317
9.6.2	Number of Discriminant Functions	318
9.6.3	Interpreting Discriminant Functions	319
9.6.3.1	Discriminant Function Plots	319
9.6.3.2	Loading Matrices	320
9.6.4	Evaluating Predictor Variables	321
9.6.5	Design Complexity: Factorial Designs	326
9.6.6	Use of Classification Procedures	327
9.6.6.1	Cross-Validation and New Cases	327
9.6.6.2	Plots	327

9.6.6.3	Jackknifed Classification	328	
9.6.6.4	Evaluating Improvement in Classification	328	
9.7	COMPARISON OF PROGRAMS		330
9.7.1	SPSS Programs	330	
9.7.2	BMD Series	333	
9.8	COMPLETE EXAMPLES OF DISCRIMINANT FUNCTION ANALYSIS		334
9.8.1	Evaluation of Assumptions	335	
9.8.1.1	Unequal Sample Sizes and Missing Data	335	
9.8.1.2	Linearity	335	
9.8.1.3	Outliers	336	
9.8.1.4	Multivariate Normality	336	
9.8.1.5	Homogeneity of Variance-Covariance Matrices	336	
9.8.1.6	Multicollinearity and Singularity	338	
9.8.2	Direct Discriminant Function Analysis	341	
9.8.3	Hierarchical Discriminant Function Analysis	353	
9.9	SOME EXAMPLES FROM THE LITERATURE		369
Chapter 10	Principal Components Analysis and Factor Analysis		372
10.1	GENERAL PURPOSE AND DESCRIPTION		372
10.2	KINDS OF RESEARCH QUESTIONS		375
10.2.1	Nature of Latent Structure	375	
10.2.2	Number of Latent Structures	376	
10.2.3	Importance of Factors and Solutions	376	
10.2.4	Testing Theory in FA	376	
10.2.5	Comparing Structure in Different Groups	376	
10.2.6	Distinguishing Among Subjects	377	
10.2.7	Estimating Scores on Latent Variables	377	
10.3	LIMITATIONS		377
10.3.1	Theoretical Issues	377	
10.3.2	Practical Issues	379	
10.3.2.1	Outliers Among Cases	379	
10.3.2.2	Sample Size	379	
10.3.2.3	Factorability of R	379	
10.3.2.4	Multicollinearity and Singularity	380	
10.3.2.5	Outliers Among Variables	380	
10.3.2.6	Outliers with Respect to the Solution	380	
10.3.2.7	Normality and Linearity	380	
10.4	FUNDAMENTAL EQUATIONS FOR ANALYSES OF STRUCTURE		381
10.5	MAJOR TYPES OF ANALYSIS OF STRUCTURE		394
10.5.1	Factor Extraction Techniques	394	
10.5.1.1	PCA Versus FA	395	
10.5.1.2	Principal Components	397	
10.5.1.3	Principal Factors	397	
10.5.1.4	Image Factor Extraction	398	
10.5.1.5	Maximum Likelihood Factor Extraction	398	
10.5.1.6	Alpha Factoring	398	

10.5.2 Rotation	399	
10.5.2.1 Orthogonal Rotation	399	
10.5.2.2 Oblique Rotation	401	
10.5.2.3 Geometric Interpretation	402	
10.5.3 Some Practical Recommendations	404	
10.6 SOME IMPORTANT ISSUES		404
10.6.1 Estimates of Communalities	404	
10.6.2 Adequacy of Extraction and Number of Factors	405	
10.6.3 Adequacy of Rotation and Simple Structure	408	
10.6.4 Importance and Internal Consistency of Factors	409	
10.6.5 Interpretations of Factors	411	
10.6.6 Factor Scores	412	
10.6.7 Comparisons Among Solutions and Groups	413	
10.7 COMPARISON OF PROGRAMS		416
10.7.1 SPSS Package	416	
10.7.2 BMD Series	418	
10.8 COMPLETE EXAMPLE OF FA		418
10.8.1 Evaluation of Limitations	419	
10.8.1.1 Outliers Among Cases	419	
10.8.1.2 Sample Size	421	
10.8.1.3 Factorability of R	421	
10.8.1.4 Multicollinearity and Singularity	421	
10.8.1.5 Outliers Among Variables	421	
10.8.1.6 Outliers with Respect to the Solution	421	
10.8.1.7 Normality and Linearity	423	
10.8.2 Principal Factors Extraction with Varimax Rotation: Comparison of Two Groups	423	
10.9 SOME EXAMPLES FROM THE LITERATURE		444
Chapter 11 An Overview of the General Linear Model		446
11.1 LINEARITY AND THE GENERAL LINEAR MODEL		446
11.2 BIVARIATE TO MULTIVARIATE STATISTICS AND OVERVIEW OF TECHNIQUES		447
11.2.1 Bivariate Form	447	
11.2.2 Simple Multivariate Form	448	
11.2.3 Full Multivariate Form	450	
11.3 ALTERNATIVE RESEARCH STRATEGIES		452
Appendix A A Skimpy Introduction to Matrix Algebra		455
A.1 ADDITION OR SUBTRACTION OF A CONSTANT TO A MATRIX		456
A.2 MULTIPLICATION OR DIVISION OF A MATRIX BY A CONSTANT		456
A.3 ADDITION AND SUBTRACTION OF TWO MATRICES		457
A.4 MULTIPLICATION, TRANSPOSES, AND SQUARE ROOTS OF MATRICES		458
A.5 MATRIX "DIVISION" (INVERSES AND DETERMINANTS)		460

A.6 EIGENVALUES AND EIGENVECTORS: PROCEDURES FOR CONSOLIDATING VARIANCE FROM A MATRIX	462
Appendix B Research Design for Complete Examples	466
Appendix C Statistical Tables	472
C.1 NORMAL CURVE AREAS	473
C.2 CRITICAL VALUES OF THE <i>t</i> DISTRIBUTION FOR $\alpha = .05$ AND $.01$, TWO-TAILED TEST	474
C.3 CRITICAL VALUES OF THE F DISTRIBUTION	475
C.4 CRITICAL VALUES OF CHI SQUARE (χ^2)	479
C.5 CRITICAL VALUES FOR SQUARED MULTIPLE CORRELATION (R^2) IN FORWARD STEPWISE SELECTION	480
C.6 CRITICAL VALUES FOR $F_{\max} (S_{\max}^2 / S_{\min}^2)$ FOR $\alpha = .05$ AND $.01$	484
C.7 PROBABILITIES FOR THE SALIENT VARIABLE SIMILARITY INDEX, <i>S</i>	486
References	489
Index	497