

Contents

Foreword	v
Preface	vii
PART I: APPROACHES TO ANALYSING DATA	1
1 Data and Statistics	3
1.1 Introduction	3
1.2 Types of Data	4
1.3 Exploratory and Confirmatory Techniques	5
1.4 The Role of Models	6
2 Mathematical and Statistical Background	9
2.1 Introduction	9
2.2 Matrices	9
2.3 Rotation of Axes	13
2.4 Some Basic Topics in Multivariate Statistics	13
2.5 Maximum Likelihood Estimation	15
2.6 Optimisation Methods	17
Exercises	20
PART II: EXPLORING MULTIVARIATE DATA	23
3 The Initial Examination of Multivariate Data	25
3.1 Introduction	25
3.2 Marginal Views of Multivariate Data	26
3.3 Andrews Plots and Chernoff Faces	33
3.4 Probability Plots	35
3.5 Rearranging Tables of Numerical Data	37
3.6 Missing Values in Multivariate Data	40
3.7 Summary	40
Exercises	41
4 Reducing the Dimensionality of Multivariate Data: Principal Components and Correspondence Analysis	45
4.1 Introduction	45
4.2 Definition of Principal Components	46
4.3 Correlations of Variables and Components	48
4.4 Rescaling Principal Components	48
4.5 Calculating Principal Component Scores	49

4.6	Choosing the Number of Components	49
4.7	A Simple Example of Principal Components Analysis	50
4.8	Numerical Examples	51
4.9	Geometrical Interpretation of Principal Components Analysis	55
4.10	Correspondence Analysis	57
4.11	Summary	62
	Exercises	63
5	Multidimensional Scaling	67
5.1	Introduction	67
5.2	Simple Examples of Multidimensional Scaling	68
5.3	Classical Multidimensional Scaling	69
5.4	Metric and Non-metric Multidimensional Scaling	73
5.5	Choosing the Number of Dimensions	76
5.6	Further Examples of the Application of Multidimensional Scaling Techniques	78
5.7	Non-Euclidean Metrics	82
5.8	Three-way Multidimensional Scaling	82
5.9	The Analysis of Asymmetric Proximity Matrices	88
5.10	A Statistical Approach to Multidimensional Scaling	91
5.11	Summary	93
	Exercises	95
6	Cluster Analysis	99
6.1	Introduction	99
6.2	Examples of the Application of Cluster Analysis	100
6.3	Agglomerative Hierarchical Clustering Techniques	101
6.4	Mixture Models for Cluster Analysis	113
6.5	Graphical Aids to Cluster Analysis	120
6.6	Summary	123
	Exercises	124
	PART III: REGRESSION MODELS	127
7	The Generalised Linear Model	129
7.1	Linear Models	129
7.2	Non-Linear Models	133
7.3	Link Functions and Error Distribution in the Generalized Linear Model	135
	Exercises	138
8	Regression and The Analysis of Variance	139
8.1	Introduction	139
8.2	Least Squares Estimation for Regression and Analysis of Variance (ANOVA) Models	139
8.3	Multiple Regression Models	140
8.4	ANOVA Models	142
8.5	Testing the Fit of a Model	145
8.6	Numerical Examples	148

8.7	Checking Assumptions by Graphical Analysis and Examination of Residuals	156
8.8	Detection of Influential Observations	162
8.9	Weighted Least Squares	163
8.10	Summary	166
	Exercises	166
9	Linear Models for Categorical Data	169
9.1	Introduction	169
9.2	Maximum Likelihood Estimation for Log-Linear and Logistic Models	169
9.3	Goodness-of-Fit Measures and the Comparison of Alternative Models	171
9.4	Some Numerical Examples	172
9.5	Model Selection Procedures	177
9.6	Linear Models for Ordinal Response Variables	182
	Exercises	189
10	Models for Rates and Survival Times	191
10.1	Introduction	191
10.2	Log-Linear Models for Rates	193
10.3	Log-Linear Hazard Models	194
10.4	Models for Grouped Survival Times	199
	Exercises	203
11	Analysis of Repeated Measures	205
11.1	Introduction	205
11.2	Markov Chains and Autoregressive Models	207
11.3	Analysis of Covariance	210
11.4	Mixed and Random Effects Models	213
11.5	Extraction of Response Features and Trend Components	216
11.6	Multivariate Tests	218
11.7	Trend Analysis for Categorical Responses	220
	Exercises	222
12	Discriminant Analysis	225
12.1	Introduction	225
12.2	Discrimination Rules	226
12.3	Methods of Discriminant Analysis	228
12.4	Selecting Variables	234
12.5	Summary	236
	Exercises	237
	PART IV: LATENT VARIABLE MODELS	239
13	Factor Analysis	241
13.1	Introduction	241
13.2	The Basic Factor Analysis Model	241
13.3	Estimating the Parameters in the Factor Model	244

13.4	Principal Factor Analysis	245
13.5	Maximum Likelihood Factor Analysis	246
13.6	Numerical Examples of PFA and MLFA	246
13.7	Determining the Number of Factors	247
13.8	Rotation of Factors	250
13.9	Confirmatory Factor Analysis	253
13.10	Factor Analysis and Principal Component Analysis Compared	254
13.11	Summary	254
	Exercises	255
14	Covariance Structure Models	257
14.1	Introduction	257
14.2	A Simple Covariance Structure Model	257
14.3	Covariance Structure Models: Some Alternative Formulations	258
14.4	Identification of Parameters in Covariance Structure Models	263
14.5	Estimating the Parameters in Covariance Structure Models	264
14.6	Assessing the Fit of a Model	265
14.7	Numerical Examples	266
14.8	Causal Models and Latent Variables – Myths and Realities	272
14.9	Summary	273
	Exercises	274
	References	277
	Appendix A Programs and Packages	291
A.1	General Purpose Packages	291
A.2	Generalised Linear Modelling	291
A.3	Specialised Programs	292
A.4	Statistical Languages	292
	Appendix B Answers to Selected Exercises	293
B.1	Chapter 2	293
B.2	Chapter 4	294
B.3	Chapter 5	295
B.4	Chapter 6	295
B.5	Chapter 7	296
B.6	Chapter 14	296
	Index	299