

Contents

List of Figures	xi
List of Tables	xv
Preface	xix
1 Methodological issues in longitudinal research	1
1.1 Introduction	1
1.2 The data box	5
1.3 The three modes of the data box: selection effects	9
1.4 Designs for longitudinal studies	19
1.5 Validity issues in longitudinal research	29
1.5.1 Statistical conclusion validity	31
1.5.2 Internal validity	31
1.5.3 External validity	35
1.5.4 Reliability and other aspects of construct validity	36
1.6 Other methodological concerns in longitudinal research	37
1.6.1 Conditional or unconditional analysis	37
1.6.2 Aggregation	40
1.6.3 Missing data	41
1.7 Further reading	44
2 Analysis of longitudinal categorical data using optimal scaling techniques	46
2.1 Introduction	46
2.1.1 Assumptions	49

2.2	Multivariate analysis using optimal scaling	50
2.2.1	Multiple Correspondence Analysis	51
2.2.2	Nonlinear Principal Component Analysis	64
2.2.3	Nonlinear Generalized Canonical Analysis	74
2.3	Longitudinal extensions to optimal scaling techniques	84
2.3.1	Flattening the data box: time points adding subjects or time points adding variables	85
2.3.2	Time as a categorical variable	98
2.3.3	Modelling time with lagged variables	110
2.4	Multivariate dynamic analysis	121
2.4.1	AR(I)MA models	121
2.4.2	The linear dynamic system	125
2.5	Multivariate dynamic analysis of categorical data	132
2.5.1	Optimal scaling extensions to linear dynamic analysis for $N = 1$	132
2.5.2	Optimal scaling extensions to linear dynamic analysis for $N > 1$	139
2.5.3	Other multivariate dynamic analysis techniques for categorical data	148
2.6	Missing data	149
2.7	Discussion	151
2.8	Further reading	154
3	Univariate and multivariate analysis of variance of longitudinal data	155
3.1	Introduction	155
3.1.1	Types of longitudinal designs	156
3.1.2	Hypothesis testing	158
3.1.3	Differences among waves	160
3.1.4	Multivariate approaches	162
3.1.5	Linear transformations of repeated measures	165
3.1.6	Structured multivariate analysis of repeated measures	166
3.1.7	Trend analysis	167
3.1.8	Relationships between univariate and structured multivariate analysis	169

3.1.9	Assumptions in univariate ANOVA of repeated measures	170
3.1.10	Practical recommendations	174
3.2	One-group/one-variable/two-waves designs	176
3.2.1	<i>t</i> -test for dependent samples	176
3.2.2	<i>t</i> -test using difference scores	177
3.2.3	ANOVA	178
3.2.4	MANOVA	179
3.2.5	Analysis of covariance	180
3.2.6	Comparison of approaches	182
3.3	Multiple-groups/one-variable/two-waves designs	183
3.3.1	Repeated measures ANOVA	183
3.3.2	MANOVA	184
3.3.3	ANOVA on difference scores	185
3.3.4	Analysis of covariance	186
3.3.5	Comparison of approaches	188
3.4	One-group/one-variable/multiple-waves designs	189
3.4.1	RB ANOVA	190
3.4.2	MANOVA	192
3.4.3	Difference scores	193
3.4.4	Analysis of covariance	194
3.5	Multiple-groups/one-variable/multiple-waves designs	194
3.6	Multiple-variables/multiple-waves designs	195
3.7	Relations with other techniques	201
3.7.1	Relations with multilevel analysis	201
3.7.2	Relations with structural equation models	202
3.8	Missing data	202
3.9	Discussion	204
3.10	Further reading	205
4	Structural equation models for longitudinal data	207
4.1	Introduction	207
4.2	Linear structural equation models	211
4.2.1	Assumptions	211
4.2.2	Structural equation and path diagrams for various models	212
4.2.3	Model specification	218

4.2.4	Model testing	221
4.2.5	Interpretation of results	231
4.3	Structural equation modelling of longitudinal data	232
4.3.1	Autoregressive models	234
4.3.2	Growth curve modelling	248
4.4	Recent developments	260
4.5	Missing data	263
4.6	Discussion	265
4.7	Further reading	268
5	Multilevel analysis of longitudinal data	269
5.1	Introduction	269
5.1.1	Multilevel data	269
5.1.2	Multilevel analysis	271
5.1.3	Multilevel analysis of longitudinal data	273
5.1.4	Software for multilevel analysis	274
5.2	A Two-Level Model	275
5.2.1	Two levels: a within-unit and a between-unit model	275
5.2.2	Single equation version: a mixed linear model	278
5.2.3	Assumptions	279
5.2.4	Estimation	280
5.3	The multilevel treatment of longitudinal data	281
5.3.1	The hierarchical interpretation of repeated measures	282
5.3.2	Two levels: a within-subject and between-subject model	283
5.3.3	A hypothetical example	285
5.3.4	The within-subject covariance structure	287
5.3.5	Hypothesis testing	288
5.3.6	Empirical examples	292
5.4	Further issues in longitudinal multilevel analysis	301
5.4.1	Time-varying covariates	301
5.4.2	Centring of explanatory variables	302
5.4.3	Other growth models	305
5.5	Properties of longitudinal multilevel analysis	306
5.5.1	Characteristics of growth	307

5.5.2	Emphasis on individual growth	308
5.5.3	Flexibility of approach	308
5.5.4	Modelling of the within-subject covariance structure	309
5.5.5	Correspondence in results	309
5.5.6	Models with more than two levels	309
5.6	Relations with (generalized) MANOVA	310
5.7	Relations with structural equation modelling	313
5.8	Missing data	314
5.9	Discussion	315
5.10	Further reading	316
6	Log-linear and Markov modelling of categorical longitudinal data	318
6.1	Introduction	318
6.1.1	Assumptions	321
6.2	Log-linear models	322
6.2.1	Notation and an example	322
6.2.2	Model specification	324
6.2.3	Log-linear analysis for a 2×2 contingency table	325
6.2.4	Log-linear analysis for more than one table	327
6.2.5	Log-linear analysis for an $I \times I$ contingency table: testing various hypotheses on change	333
6.3	Markov models	340
6.3.1	Simple Markov models	343
6.3.2	Latent Markov models	349
6.3.3	Mixed Markov models	359
6.3.4	Latent mixed Markov models	362
6.4	Log-linear analysis and Markov models	364
6.5	Missing data	366
6.6	Discussion	367
6.7	Further reading	368
7	Epilogue	371
7.1	Relations between longitudinal analysis techniques	373
7.2	Choosing a longitudinal analysis method	381
7.3	Longitudinal analysis techniques we did not discuss	383

References	385
Notes on contributors	407
Author index	410
Subject index	416

List of Figures

1.1	Basic Data Relation Matrix (BDRM) or Data Box Proposed by Cattell (1946, 1952, 1988)	6
1.2	Six Correlational or Relational Matrices and Techniques Obtainable from the Three Facets Persons, Variables and Occasions (After Cattell, 1988)	7
1.3	Positive Correlation Computed over Time Points	32
1.4	Negative Correlation Computed over Time Points	33
2.1	Nominal $N \times 3$ Data Matrix, Corresponding Indicator Matrices, and Quantifications of the First Variable	53
2.2	Artificial Data: HOMALS Discrimination Measures	59
2.3	Artificial Data: HOMALS Category Quantifications	62
2.4	Artificial Data: HOMALS Object Scores	63
2.5	Artificial Data: PRINCALS Multiple and Single Category Coordinates of Educational Level and Age	70
2.6	Artificial Data: PRINCALS Category Quantifications of Multiple Nominal Variables and Single Category Coordinates of Single Variables	72
2.7	Artificial Data: PRINCALS Object Scores	73
2.8	Artificial Data: OVERALS Component Loadings of Single Variables	80
2.9	Artificial Data: OVERALS Category Centroids and Connected Projected Centroids	82
2.10	Artificial Data: OVERALS Object Scores	83
2.11	Slicing and Vertical Stacking of the Data Box	86
2.12	Slicing and Horizontal Stacking of the Data Box	90

2.13	Care Substitution Data: Category Quantifications of the HOMALS Solution	95
2.14	Care Substitution Data: Average Object Scores of Indication Subgroups in Experimental and Control Group at the Measurement Occasions	97
2.15	Categorical Time Variable Set-up	99
2.16	Attachment Data: Design of Categorical Time Variable OVERALS Analysis	105
2.17	Attachment Data: Component Loadings of Quantified Mother and Child Variables	107
2.18	Attachment Data: Developments for Three Attachment Subgroups	109
2.19	Relating Lag(1) Variable to Lag(0) Variable	111
2.20	Relating Lag(1) Variable to its Lag(0) Version	112
2.21	Lag(7) and Lag(1) Versions of Variables Relating to Lag(0) Variables	113
2.22	Relating Lag(1) and Lag(0) Versions of Variables for N Subjects	114
2.23	Attachment Data: Lagged Analysis Design	118
2.24	Attachment Data: Component Loadings of Quantified Mother and Child Variables in Lagged Analysis	120
2.25	Autoregressive Model	122
2.26	Moving Average Model	124
2.27	Linear Dynamic System	127
2.28	Linear Dynamic System with Transition Matrices, Disturbances and Error Terms	128
2.29	Shift Matrix S	130
2.30	Blood Pressure Data: Relation of WEEKDAY and Lag(1) Version of MEDICATION with Diastolic and Systolic Blood Pressure	137
2.31	Menstrual Distress Data: Synchronized Latent State Scores for Subgroups	145
2.32	Deletion of Records in Lagged Analysis Because of Missing Occasion	152
4.1	Simple Regression Model	213
4.2	Multiple Regression Model	214
4.3	Exploratory Factor Analysis Model	215

4.4	Confirmatory Factor Analysis Model	216
4.5	MIMIC Model	217
4.6	Hypothetical Model	219
4.7	Simple Measurement Model	223
4.8	Simple Measurement Model	226
4.9	Longitudinal Model with Correlated Error Terms	233
4.10	Markov Model for Observed Variables	234
4.11	Markov Model for Latent Variables	236
4.12	Dynamic Factor Model with Predictor Variables	237
4.13	Depression Data: Linear Dynamic Model	242
4.14	Depression Data: Linear Dynamic Model with Phantom Variables	244
4.15	Depression Data: Results for Linear Dynamic Model	246
4.16	Growth Curve Model of Latent Ability	249
4.17	Depression Data: Growth Curve Model with Demographic Predictors	253
4.18	Depression Data: Growth Curve Results with Demographic Predictors	254
4.19	Depression Data: Growth Curve Model with Demographic and Time-Varying Predictors	257
4.20	Depression Data: Growth Curve Results with Demographic and Time-Varying Predictors	258
5.1	Dental Data, Predicted Growth Curves	295
5.2	Rat Data, Predicted Growth Curves for 10 Rats	298
6.1	Overview of Markov Models	342
6.2	Simple Markov Model for Three Time Points	344
6.3	Example of a Transition Matrix F	345
6.4	Latent Markov Model for Three Time Points, One Observed Variable	350
6.5	Example of a Matrix H_t	351
6.6	Latent Markov Model for Three Time Points, Two Observed Variables	353
6.7	Latent Markov Model for Three Time Points, Two Observed Variables and Two Latent Variables Set Equal	354

6.8	Example of Transition Matrix for Stage-Sequential Model	369
7.1	Paths for Choice of Analysis Method	382

List of Tables

1.1	Schematic Representation of the Simultaneous Cross-Sectional Design (After Keeves, 1997)	22
1.2	Schematic Representation of a Trend Study	23
1.3	Schematic Representation of a Time Series Study	23
1.4	Schematic Representation of Intervention Studies	28
1.5	Schematic Representation of a Longitudinal Panel Study	29
2.1	Artificial Data: Subject Number, Raw Data, Variables and Categories	57
2.2	Artificial Data: HOMALS Eigenvalues and Discrimination Measures	58
2.3	Artificial Data: Variables, Categories, Frequencies and HOMALS Category Quantifications	60
2.4	Artificial Data: Variables and Measurement Levels used in PRINCALS Analysis	69
2.5	Artificial Data: Sets, Variables and Measurement Levels used in OVERALS Analysis	78
2.6	Artificial Data: Summary of OVERALS Solution: Loss per Set, Mean Loss, Fit and Eigenvalues	79
2.7	Care Substitution Data: Well-being and Care Variables	94
2.8	Attachment Data: Mother and Child Behaviour Variables	104
2.9	Blood Pressure Data: Correlations of Quantified Input Variables and Output Variables with the Latent State	137

2.10	Blood Pressure Data: Category Quantifications of Medication	138
2.11	Menstrual Distress Data: Concentration Complaints and Menstrual Distress Variables	142
2.12	Menstrual Distress Data: Correlations of Quantified Input Variables and Concentration Variable with the Latent State	143
2.13	Quantification of Missing Data	150
3.1	Example of Split Plot Design with Data Vector y and column c_1 through c_8 of the Matrix C of Contrast Vectors	163
3.2	Analysis of Variance of the Data of Table 3.1	164
3.3	Example of Split Plot Design with Data Matrix Y and Matrices H and C of Contrast Vectors	167
3.4	ANOVAs and MANOVA of the Transformed Data of Table 3.3	168
3.5	Example of a One-Group/One-Variable/Two-Waves Data Matrix	177
3.6	ANOVA of One-Group/One-Variable/Two-Waves Data as an RB-2 Design	179
3.7	ANCOVA via Regression Analysis of One-Group/One-Variable/ Two-Waves data	182
3.8	Example of a Two-Group/One-Variable/Two-Waves Data Matrix	183
3.9	ANOVA of Two-Group/One-Variable/Two-Waves Data as an SPF 2.2 Design	184
3.10	Tests of the Homogeneity of the Within-groups Variances and Covariances	185
3.11	ANOVA of Two-Group/Two-Wave Difference Scores as a CRF 2 Design	186
3.12	ANCOVA of Two-Group/One-Variable/Two-Waves Data as a CR 2 Design	187
3.13	Example of a One-Group/One-Variable/Three-Waves Data Matrix	191
3.14	ANOVA of One-Group/One-Variable/Three-Waves Data as an RB-3 Design	191

3.15	MANOVA of Two-Groups/One-Variable/Three-Waves Data	196
3.16	Data Matrix of Two-Groups/Three-Variables/Three-Waves Design	197
3.17	Transformation Matrix for Three-Variables/Three-Waves Data Matrix	198
3.18	MANOVA of Two-Groups/Three-Variable/Three-Waves Data	199
4.1	Depression Data: Variables in the Analysis	241
5.1	Dental Data, Fixed Parameter and Variance Component Estimates, Standard Errors and t -ratios for Linear Growth Model	293
5.2	Rat Data, Fixed Parameter and Variance Component Estimates, Standard Errors and t -ratios for Linear Growth Model	296
5.3	School Data, Results for Quadratic Growth Model with Four Different Structures for Σ	300
6.1	Example Data: Employment Status 1984 and 1985	323
6.2	Example Data: Employment Status 1984 and Mobility 1985	323
6.3	Danish and British Social Mobility Data	328
6.4	Danish and British Social Mobility Data: Summary of Test Results	331
6.5	Danish and British Social Mobility Data: Standardized Residuals	332
6.6	Danish Social Mobility Data	333
6.7	Danish Social Mobility Data: Standardized Residuals for the Independence Model	334
6.8	Danish Social Mobility Data: Standardized Residuals for the Quasi-Independence Model	335
6.9	Danish Social Mobility Data: Standardized Residuals for the Symmetry Model	337
6.10	Danish Social Mobility Data: Standardized Residuals for the Quasi-Symmetry Model	338
6.11	Danish Social Mobility Data: Summary Table of Goodness-of-Fit Tests for Respective Models	339

6.12	Los Angeles Panel Data for State of Depression at Four Consecutive Time Points	343
6.13	Los Angeles Panel Data: Results for the Simple Markov Model	348
6.14	Los Angeles Panel Data: Results for the Latent Markov Model	355
6.15	Los Angeles Panel Data: Results for the Latent Class Model	357
6.16	Los Angeles Panel Data: Results for the Mixed Markov Model	360
6.17	Los Angeles Panel Data: Results for the Latent Mixed Markov Model	363
6.18	Los Angeles Panel Data: Turnover Tables	364
6.19	Los Angeles Panel Data: Observed Transition Proportions	365
7.1	Characteristics of Longitudinal Analysis Techniques	374