

2<sup>nd</sup> Edition

# MULTILEVEL ANALYSIS

An  
Introduction to  
Basic and Advanced  
Multilevel Modeling

Tom A B **SNIJDERS**  
Roel J **BOSKER**



# Contents

<b>Preface to the Second Edition</b>	<b>x</b>
<b>Preface to the First Edition</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Multilevel analysis	1
1.1.1 Probability models	2
1.2 This book	3
1.2.1 Prerequisites	5
1.2.2 Notation	5
<b>2 Multilevel Theories, Multistage Sampling, and Multilevel Models</b>	<b>6</b>
2.1 Dependence as a nuisance	6
2.2 Dependence as an interesting phenomenon	8
2.3 Macro-level, micro-level, and cross-level relations	9
2.4 Glommary	13
<b>3 Statistical Treatment of Clustered Data</b>	<b>14</b>
3.1 Aggregation	14
3.2 Disaggregation	16
3.3 The intraclass correlation	17
3.3.1 Within-group and between-group variance	19
3.3.2 Testing for group differences	22
3.4 Design effects in two-stage samples	23
3.5 Reliability of aggregated variables	25
3.6 Within- and between-group relations	26
3.6.1 Regressions	27
3.6.2 Correlations	32
3.6.3 Estimation of within- and between-group correlations	34
3.7 Combination of within-group evidence	36
3.8 Glommary	39
<b>4 The Random Intercept Model</b>	<b>41</b>
4.1 Terminology and notation	42
4.2 A regression model: fixed effects only	43

4.3	Variable intercepts: fixed or random parameters?	44
4.3.1	When to use random coefficient models	46
4.4	Definition of the random intercept model	49
4.5	More explanatory variables	54
4.6	Within- and between-group regressions	56
4.7	Parameter estimation	60
4.8	'Estimating' random group effects: posterior means	62
4.8.1	Posterior confidence intervals	64
4.9	Three-level random intercept models	67
4.10	Glommary	71
<b>5</b>	<b>The Hierarchical Linear Model</b>	<b>74</b>
5.1	Random slopes	74
5.1.1	Heteroscedasticity	75
5.1.2	Do not force $\tau_{01}$ to be 0!	76
5.1.3	Interpretation of random slope variances	77
5.2	Explanation of random intercepts and slopes	80
5.2.1	Cross-level interaction effects	81
5.2.2	A general formulation of fixed and random parts	86
5.3	Specification of random slope models	87
5.3.1	Centering variables with random slopes?	87
5.4	Estimation	89
5.5	Three or more levels	90
5.6	Glommary	92
<b>6</b>	<b>Testing and Model Specification</b>	<b>94</b>
6.1	Tests for fixed parameters	94
6.1.1	Multiparameter tests for fixed effects	96
6.2	Deviance tests	97
6.2.1	More powerful tests for variance parameters	98
6.3	Other tests for parameters in the random part	100
6.3.1	Confidence intervals for parameters in the random part	100
6.4	Model specification	102
6.4.1	Working upward from level one	104
6.4.2	Joint consideration of level-one and level-two variables	106
6.4.3	Concluding remarks on model specification	107
6.5	Glommary	108
<b>7</b>	<b>How Much Does the Model Explain?</b>	<b>109</b>
7.1	Explained variance	109
7.1.1	Negative values of $R^2$ ?	110
7.1.2	Definitions of the proportion of explained variance in two-level models	111
7.1.3	Explained variance in three-level models	113
7.1.4	Explained variance in models with random slopes	113

7.2	Components of variance	114
7.2.1	Random intercept models	114
7.2.2	Random slope models	116
7.3	Glommary	117
<b>8</b>	<b>Heteroscedasticity</b>	<b>119</b>
8.1	Heteroscedasticity at level one	119
8.1.1	Linear variance functions	119
8.1.2	Quadratic variance functions	123
8.2	Heteroscedasticity at level two	128
8.3	Glommary	129
<b>9</b>	<b>Missing Data</b>	<b>130</b>
9.1	General issues for missing data	131
9.1.1	Implications for design	133
9.2	Missing values of the dependent variable	133
9.3	Full maximum likelihood	134
9.4	Imputation	135
9.4.1	The imputation method	137
9.4.2	Putting together the multiple results	140
9.5	Multiple imputations by chained equations	144
9.6	Choice of the imputation model	148
9.7	Glommary	149
<b>10</b>	<b>Assumptions of the Hierarchical Linear Model</b>	<b>152</b>
10.1	Assumptions of the hierarchical linear model	153
10.2	Following the logic of the hierarchical linear model	154
10.2.1	Include contextual effects	154
10.2.2	Check whether variables have random effects	155
10.2.3	Explained variance	156
10.3	Specification of the fixed part	156
10.4	Specification of the random part	158
10.4.1	Testing for heteroscedasticity	159
10.4.2	What to do in case of heteroscedasticity	161
10.5	Inspection of level-one residuals	161
10.6	Residuals at level two	165
10.7	Influence of level-two units	167
10.8	More general distributional assumptions	172
10.9	Glommary	173
<b>11</b>	<b>Designing Multilevel Studies</b>	<b>176</b>
11.1	Some introductory notes on power	177
11.2	Estimating a population mean	179
11.3	Measurement of subjects	180
11.4	Estimating association between variables	180
11.4.1	Cross-level interaction effects	185
11.5	Allocating treatment to groups or individuals	187

11.6	Exploring the variance structure	188
11.6.1	The intraclass correlation	188
11.6.2	Variance parameters	190
11.7	Glommary	191
<b>12</b>	<b>Other Methods and Models</b>	<b>194</b>
12.1	Bayesian inference	194
12.2	Sandwich estimators for standard errors	197
12.3	Latent class models	201
12.4	Glommary	203
<b>13</b>	<b>Imperfect Hierarchies</b>	<b>205</b>
13.1	A two-level model with a crossed random factor	206
13.2	Crossed random effects in three-level models	210
13.3	Multiple membership models	210
13.4	Multiple membership multiple classification models	213
13.5	Glommary	215
<b>14</b>	<b>Survey Weights</b>	<b>216</b>
14.1	Model-based and design-based inference	217
14.1.1	Descriptive and analytic use of surveys	217
14.2	Two kinds of weights	219
14.3	Choosing between model-based and design-based analysis	222
14.3.1	Inclusion probabilities and two-level weights	223
14.3.2	Exploring the informativeness of the sampling design	225
14.4	Example: Metacognitive strategies as measured in the PISA study	231
14.4.1	Sampling design	231
14.4.2	Model-based analysis of data divided into parts	233
14.4.3	Inclusion of weights in the model	235
14.5	How to assign weights in multilevel models	236
14.6	Appendix. Matrix expressions for the single-level estimators	244
14.7	Glommary	244
<b>15</b>	<b>Longitudinal Data</b>	<b>247</b>
15.1	Fixed occasions	248
15.1.1	The compound symmetry model	249
15.1.2	Random slopes	253
15.1.3	The fully multivariate model	255
15.1.4	Multivariate regression analysis	260
15.1.5	Explained variance	261
15.2	Variable occasion designs	263
15.2.1	Populations of curves	263
15.2.2	Random functions	263
15.2.3	Explaining the functions	274
15.2.4	Changing covariates	276
15.3	Autocorrelated residuals	280
15.4	Glommary	280

<b>16 Multivariate Multilevel Models</b>	<b>282</b>
16.1 Why analyze multiple dependent variables simultaneously?	283
16.2 The multivariate random intercept model	283
16.3 Multivariate random slope models	288
16.4 Glommary	288
<b>17 Discrete Dependent Variables</b>	<b>289</b>
17.1 Hierarchical generalized linear models	289
17.2 Introduction to multilevel logistic regression	290
17.2.1 Heterogeneous proportions	290
17.2.2 The logit function: Log-odds	293
17.2.3 The empty model	295
17.2.4 The random intercept model	297
17.2.5 Estimation	300
17.2.6 Aggregation	301
17.3 Further topics on multilevel logistic regression	302
17.3.1 Random slope model	302
17.3.2 Representation as a threshold model	303
17.3.3 Residual intraclass correlation coefficient	304
17.3.4 Explained variance	305
17.3.5 Consequences of adding effects to the model	307
17.4 Ordered categorical variables	310
17.5 Multilevel event history analysis	313
17.6 Multilevel Poisson regression	314
17.7 Glommary	320
<b>18 Software</b>	<b>323</b>
18.1 Special software for multilevel modeling	323
18.1.1 HLM	324
18.1.2 MLwiN	325
18.1.3 The MIXOR suite and SuperMix	325
18.2 Modules in general-purpose software packages	327
18.2.1 SAS procedures VARCOMP, MIXED, GLIMMIX, and NL MIXED	327
18.2.2 R	328
18.2.3 Stata	328
18.2.4 SPSS, commands VARCOMP and MIXED	329
18.3 Other multilevel software	329
18.3.1 PinT	329
18.3.2 Optimal Design	330
18.3.3 MLPowSim	330
18.3.4 Mplus	330
18.3.5 Latent Gold	330
18.3.6 REALCOM	330
18.3.7 WinBUGS	331
<b>References</b>	<b>332</b>
<b>Index</b>	<b>348</b>