Contents

Preface **xxi**

1 Linear variance-components models **1**

1.1 Introduction . . . 1
1.2 How reliable are expiratory flow measurements? . . . 1
1.3 The variance-components model . . . 4
1.3.1 Model specification and path diagram . . . 4
1.3.2 Error components, variance components, and reliability . . . 7
1.3.3 Intraclass correlation . . . 8
1.4 Modeling the Mini Wright measurements . . . 8
1.4.1 Estimation using xtreg . . . 9
1.4.2 Estimation using xtmixed . . . 10
1.4.3 Estimation using gllamm . . . 11
1.4.4 Relative and absolute agreement . . . 14
1.5 Estimation methods . . . 15
1.6 Assigning values to the random intercepts . . . 16
1.6.1 Maximum likelihood estimation . . . 17
Implementation via OLS regression . . . 17
Implementation via the mean total residual . . . 19
1.6.2 Empirical Baycs prediction . . . 19
1.6.3 ∇ Empirical Bayes variances . . . 24
1.7 Summary and further reading . . . **25**
1.8 Exercises . . . 26

2 Linear random-intercept models **31**

2.1 Introduction . . . 31
2.2 Are tax preparers useful? . . . 31
2.3 The longitudinal data structure . . . 32
2.4 Panel data and correlated residuals . . . 32
2.5 The random-intercept model . . . 34
2.5.1 Estimation using xtreg . . . 36
2.5.2 Estimation using xtmixed . . . 38
2.6 Different kinds of effects in panel models . . . 39
2.6.1 Between-taxpayer effects . . . 39
2.6.2 Within-taxpayor effects . . . 40
2.6.3 Relations among the estimators . . . 42
2.7 Endogeneity and between-taxpayer effects . . . 42
2.8 Residual diagnostics . . . 45
2.9 Summary and further reading . . . 50
2.10 Exercises . . . 51

3 Linear random-coefficient and growth-curve models **57**

3.1 Introduction . . . 57
3.2 How effective are different schools? . . . 57
3.3 Separate linear regressions for each school . . . 58
3.4 The random-coefficient model . . . 61
3.4.1 Specification and interpretation of a random-coefficient model . . . 61
3.4.2 Estimation arid prediction using xtmixed . . . 68
Estimation of random-intercept model . . . 68
Estimation of random-coefficient model . . . 69
Empirical Bayes prediction using xtmixed . . . 72
3.4.3 Estimation and prediction using gilamm . . . 74
Estimation of random-intercept model . . . 75
Estimation of random-coefficient model . . . 75
Empirical Bayes prediction . . . 77

3.5 How do children grow? . . . 78
3.6 Growth-curve modeling . . . 79
3.6.1 Observed growth trajectories . . . 79
3.6.2 Estimation using xtmixed . . . 80
Quadratic growth model with random intercept . . . 80
Quadratic growth model with random intercept and random slope 80
Including a child-level covariate . . . 83
3.6.3 Estimation using gllamm . . . 84
Quadratic growth model with random intercept . . . 84
Quadratic growth model with random intercept and random slope 85
Including a child-level covariate . . . 86
3.7 Two-stage model formulation . . . 87
3.7.1 Model specification . . . 88
3.7.2 Estimation . . . 88
3.8 Prediction of trajectories for individual children . . . 91
3.9 Complex level-1 variation or heteroskedasticity . . . 92
3.10 Summary and further reading . . . 94
3.11 Exercises . . . 95

4 Dichotomous or binary responses 101

4.1 Models for dichotomous responses . . . 101
4.1.1 Generalized linear model formulation . . . 101
4.1.2 Latent-response formulation . . . 107
Logistic regression . . . 108
Probit regression . . . 109
4.2 Which treatment is best for toenail infection? . . . 1)1
4.3 The longitudinal data structure . . . 112
4.4 Population-averaged or marginal probabilities . . . 113
4.5 Random-intercept logistic regression . . . 116
4.6 Subject-specific vs. population-averaged relationships . . . 120

4.7 Maximum likelihood estimation using adaptive quadrature 124
4.7.1 Some practical considerations 128
4.8 Empirical Bayes (EB) predictions 129
4.8.1 EB prediction of random effects 129
4.8.2 EB prediction of response probabilities 130
4.9 Other approaches to clustered dichotomous data 131
4.9.1 Conditional logistic regression 131
4.9.2 Generalized estimating equations (GEE) 132
4.10 Summary and further reading 134
4.11 Exercises 134

5 Ordinal responses **143**
5.1 Introduction 143
5.2 Cumulative models for ordinal responses 143
5.2.1 Generalized linear model formulation 143
5.2.2 Latent-response formulation 144
5.2.3 Proportional odds 147
5.2.4 Identification 148
5.3 Are antipsychotic drugs effective for patients with schizophrenia? . . . 149
5.4 Longitudinal data structure and graphs 150
5.4.1 The longitudinal data structure 150
5.4.2 Plotting cumulative proportion.[1]) 151
5.4.3 Plotting cumulative logits and transforming the time scale 152
5.5 A proportional-odds model 154
5.5.1 Model specification 154
5.5.2 Estimation 155
5.0 A random-intercept proportional-odds model 157
5.6.1 Model specification 157
5.6.2 Estimation 158
5.7 A random-coefficient proportional-odds model 159

5.7.1 Model specification . . . 159
5.7.2 Estimation . . . 159
5.8 Marginal and patient-specific probabilities . . . 161
5.8.1 Marginal probabilities . . . 161
5.8.2 Patient-specific cumulative response probabilities . . . 163
5.9 Do experts differ in their grading of student essays? . . . 165
5.10 A random-intercept model with grader bias . . . 165
5.10.1 Model specification . . . 165
5.10.2 Estimation . . . 166
5.11 * Including grader-specific measurement error variances . . . 167
5.11.1 Model specification . . . 167
5.11.2 Estimation . . . 167
5.12 * Including grader-specific thresholds . . . 169
5.12.1 Model specification . . . 169
5.12.2 Estimation . . . 170
5.13 Summary and further reading . . . 175
5.14 Exercises . . . 175

6 Counts 181

6.1 Introduction . . . 181
6.2 Types of counts . . . 181
6.3 Poisson models for counts . . . 182
6.4 Did the German health-care reform reduce the number of doctor visits? 184
6.5 Longitudinal data structure . . . 184
6.6 Poisson regression ignoring overdispersion and clustering . . . 185
6.6.1 Model specification . . . 185
6.6.2 Estimation . . . 185
6.7 Poisson regression with overdispersion but ignoring clustering . . . 187
6.7.1 Using a level-1 random intercept . . . 187
Model specification . . . 187
Estimation . . . 187

G.7.2 Quasilikelihood . . . 188
Specification . . . 188
Estimation . . . 189
6.8 Random-intercept Poisson regression . . . 190
6.8.1 Model specification . . . 190
6.8.2 Estimation . . . 191
6.9 Random-coefficient Poisson regression . . . 193
6.9.1 Model specification . . . 193
6.9.2 Estimation . . . 194
6.10 Other approaches to clustered counts . . . 196
6.10.1 Conditional Poisson regression . . . 196
6.10.2 Generalized estimating equations (GEE) . . . 197
6.11 Which Scottish counties have **a** high risk of lip cancer? . . . 197
6.12 Standardized mortality ratios . . . 198
6.13 Random-intercept Poisson regression . . . 201
6.13.1 Model specification . . . 201
6.13.2 Estimation . . . 201
6.13.3 Introducing a county-level covariate . . . 202
6.13.4 Prediction . . . 203
6.14 ♦ Nonparametric maximum likelihood estimation . . . 206
6.14.1 Specification . . . 206
6.14.2 Estimation . . . 206
6.14.3 Prediction . . . 210
6.15 Summary and further reading . . . 211
6.16 Exercises . . . 212

7 Higher level models and nested random effects 217

7.1. Introduction . . . 217
7.2 Which method is best for measuring expiratory flow? . . . 218
7.3 Two-level variance-components models . . . 219

7.3.1 Model specification . . . 219
7.3.2 Estimation . . . 219
7.4 Three-level variance-components models . . . 222
7.4.1 Model specification . . . 222
7.4.2 Different types of intraclass correlation . . . 224
7.4.3 Three-stage formulation . . . 225
7.4.4 Estimation using xtmixed . . . 225
7.4.5 Prediction using xtmixed . . . 228
7.5 Did the Guatemalan immunization campaign work? . . . 229
7.6 A three-level logistic random-intercept model . . . 230
7.6.1 Model specification . . . 230
7.6.2 Different types of intraclass correlations for the latent responses . 230
7.6.3 Three-stage formulation . . . 231
7.6.4 Estimation . . . 231
7.6.5 Introducing a random coefficient at level 3 . . . 235
7.6.6 Prediction . . . 239
7.7 Summary and further reading . . . 240
7.8 Exercises . . . 241

8 Crossed random effects **249**

8.1 Introduction . . . 249
8.2 How does investment depend on expected profit and capital stock? . . . 250
8.3 A two-way error-components model . . . 251
8.3.1 Model specification . . . 251
8.3.2 Intraclass correlations . . . 252
8.3.3 Estimation . . . 252
8.3.4 Prediction . . . 254
8.4 How much do primary and secondary schools affect attainment at age 16? . . . 257
8.5 An additive crossed random-effects model . . . 259
8.5.1 Specification . . . 259

8.5.2 Estimation . . . 259
8.6 Including a random interaction . . . 260
8.6.1 Model specification . . . 260
8.6.2 Intraclass correlations . . . 261
8.6.3 Estimation . . . 262
8.6.4 Some diagnostics . . . 264
8.7 * A trick requiring fewer random effects . . . 266
8.8 Summary and further reading . . . 269
8.9 Exercises . . . 269

A Syntax for gllamm, eq, and gllapred **275**

B Syntax for gllamm **281**

C Syntax for gllapred **293**

D Syntax for gllasim **297**

References **301**

Author index **309**

Subject index **313**