Contents

Preface ix

Dedication xi

I Methodology xiii

1 The omni-presence of latent variables 1
- 1.1 Introduction 1
- 1.2 'True' variable measured with error 2
- 1.3 Hypothetical constructs 4
- 1.4 Unobserved heterogeneity 9
- 1.5 Missing values and counterfactuals 14
- 1.6 Latent responses 16
- 1.7 Generating flexible distributions 17
- 1.8 Combining information 18
- 1.9 Summary 18

2 Modeling different response processes 21
- 2.1 Introduction 21
- 2.2 Generalized linear models 22
- 2.3 Extensions of generalized linear models 27
- 2.4 Latent response formulation 33
- 2.5 Modeling durations or survival 39
- 2.6 Summary and further reading 46

3 Classical latent variable models 49
- 3.1 Introduction 49
- 3.2 Multilevel regression models 49
- 3.3 Factor models and item response models 63
- 3.4 Latent class models 74
- 3.5 Structural equation models with latent variables 75
- 3.6 Longitudinal models 80
- 3.7 Summary and further reading 93

4 General model framework 95
- 4.1 Introduction 95

4.2 Response model 96
4.3 Structural model for the latent variables 108
4.4 Distribution of the disturbances 112
4.5 Parameter restrictions and fundamental parameters 117
4.6 Reduced form of the latent variables and linear predictor 120
4.7 Moment structure of the latent variables 121
4.8 Marginal moment structure of observed and latent responses 123
4.9 Reduced form distribution and likelihood 127
4.10 Reduced form parameters 130
4.11 Summary and further reading 132

5 Identification and equivalence 135
5.1 Introduction 135
5.2 Identification 136
5.3 Equivalence 152
5.4 Summary and further reading 157

6 Estimation 159
6.1 Introduction 159
6.2 Maximum likelihood: Closed form marginal likelihood 160
6.3 Maximum likelihood: Approximate marginal likelihood 161
6.4 Maximizing the likelihood 174
6.5 Nonparametric maximum likelihood estimation 182
6.6 Restricted/Residual maximum likelihood (REML) 184
6.7 Limited information methods 185
6.8 Maximum quasi-likelihood 191
6.9 Generalized Estimating Equations (GEE) 197
6.10 Fixed effects methods 200
6.11 Bayesian methods 204
6.12 Summary 214
Appendix: Some software and references 216

7 Assigning values to latent variables 221
7.1 Introduction 221
7.2 Posterior distributions 222
7.3 Empirical Bayes (EB) 225
7.4 Empirical Bayes modal (EBM) 235
7.5 Maximum likelihood 237
7.6 Relating the scoring methods in the 'linear case' 241
7.7 Ad hoc scoring methods 241
7.8 Some uses of latent scoring and classification 243
7.9 Summary and further reading 247
Appendix: Some software 249

8 Model specification and inference **251**
8.1 Introduction 251
8.2 Statistical modeling 251
8.3 Inference (likelihood based) 257
8.4 Model selection: Relative fit criteria 262
8.5 Model adequacy: Global absolute fit criteria 267
8.6 Model diagnostics: Local absolute fit criteria 272
8.7 Summary and further reading 280

II Applications **283**

9 Dichotomous responses **285**
9.1 Introduction 285
9.2 Respiratory infection in children: A random intercept model 285
9.3 Diagnosis of myocardial infarction: A latent class model 288
9.4 Arithmetic reasoning: Item response models 292
9.5 Nicotine gum and smoking cessation: A meta-analysis 299
9.6 Wives' employment transitions: Markov models with unobserved heterogeneity 307
9.7 Counting snowshoe hares: Capture-recapture models with heterogeneity 311
9.8 Attitudes to abortion: A multilevel item response model 315
9.9 Summary and further reading 318

10 Ordinal responses **321**
10.1 Introduction 321
10.2 Cluster randomized trial of sex education: Latent growth curve model 321
10.3 Political efficacy: Factor dimensionality and item-bias 326
10.4 Life satisfaction: Ordinal scaled probit factor models 338
10.5 Summary and further reading 347

11 Counts **349**
11.1 Introduction 349
11.2 Prevention of faulty teeth in children: Modeling overdispersion 349
11.3 Treatment of epilepsy: A random coefficient model 355
11.4 Lip cancer in Scotland: Disease mapping 361
11.5 Summary and further reading 370

12 Durations and survival **373**
12.1 Introduction 373
12.2 Modeling multiple events clustered duration data 373
12.3 Onset of smoking: Discrete time frailty models 375
12.4 Exercise and angina: Proportional hazards random effects and factor models 382

12.5 Summary and further reading 390

13 Comparative responses **393**
13.1 Introduction 393
13.2 Heterogeneity and 'Independence from Irrelevant Alternatives' 393
13.3 Model structure 395
13.4 British general elections: Multilevel models for discrete choice and rankings 397
13.5 Post-materialism: A latent class model for rankings 404
13.6 Consumer preferences for coffee makers: A conjoint choice model 408
13.7 Summary and further reading 415

14 Multiple processes and mixed responses **417**
14.1 Introduction 417
14.2 Diet and heart disease: A covariate measurement error model 417
14.3 Herpes and cervical cancer: A latent class covariate measurement error model for a case-control study 422
14.4 Job training and depression: A complier average causal effect model 427
14.5 Physician advice and drinking: An endogenous treatment model 432
14.6 Treatment of liver cirrhosis: A joint survival and marker model 438
14.7 Summary and further reading 443

References **445**

Index **487**

Author index **496**