

Contents

Symbols and Acronyms	xxi
1 • <u>Introduction to Measurement</u>	1
Measurement / 1	
Some Measurement Issues / 3	
Item Response Theory / 5	
Classical Test Theory / 5	
Latent Class Analysis / 8	
Summary / 9	
2 • <u>The One-Parameter Model</u>	12
Conceptual Development of the Rasch Model / 12	
The One-Parameter Model / 17	
The One-Parameter Logistic Model and the Rasch Model / 20	
Assumptions Underlying the Model / 21	
An Empirical Data Set: The Mathematics Data Set / 23	
Conceptually Estimating an Individual's Location / 23	
Some Pragmatic Characteristics of Maximum Likelihood Estimates / 28	
The Standard Error of Estimate and Information / 29	
An Instrument's Estimation Capacity / 32	
Summary / 35	
3 • <u>Joint Maximum Likelihood Parameter Estimation</u>	42
Joint Maximum Likelihood Estimation / 42	
Indeterminacy of Parameter Estimates / 44	
How Large a Calibration Sample? / 45	
<i>Example: Application of the Rasch Model to the Mathematics Data, JMLE,</i>	
<i>BIGSTEPS / 46</i>	
<i>Example: Application of the Rasch Model to the Mathematics Data, JMLE,</i>	
<i>mixRasch / 68</i>	
Validity Evidence / 75	

Summary of the Application of the Rasch Model / 76
Summary / 77

4 • Marginal Maximum Likelihood Parameter Estimation 86

Marginal Maximum Likelihood Estimation / 86
Estimating an Individual's Location: Expected A Posteriori / 93
 Example: Application of the Rasch Model to the Mathematics Data, MMLE, BILOG-MG / 98
Metric Transformation and the Total Characteristic Function / 111
 Example: Application of the Rasch Model to the Mathematics Data, MMLE, mirt / 115
Summary / 125

5 • The Two-Parameter Model 135

Conceptual Development of the Two-Parameter Model / 135
Information for the Two-Parameter Model / 137
Conceptual Parameter Estimation for the 2PL Model / 139
How Large a Calibration Sample? / 140
Metric Transformation, 2PL Model / 142
 Example: Application of the 2PL Model to the Mathematics Data, MMLE, BILOG-MG / 143
Fit Assessment: An Alternative Approach for Assessing Invariance / 146
 Example: Application of the 2PL Model to the Mathematics Data, MMLE, mirt / 152
Information and Relative Efficiency / 162
Summary / 165

6 • The Three-Parameter Model 179

Conceptual Development of the Three-Parameter Model / 179
Additional Comments about the Pseudo-Guessing Parameter, χ_j / 182
Conceptual Parameter Estimation for the 3PL Model / 183
How Large a Calibration Sample? / 187
Assessing Conditional Independence / 188
 Example: Application of the 3PL Model to the Mathematics Data, MMLE, BILOG-MG / 192
Fit Assessment: Conditional Independence Assessment / 195
Fit Assessment: Model Comparison / 198
 Example: Application of the 3PL Model to the Mathematics Data, MMLE, mirt / 200
Assessing Person Fit: Appropriateness Measurement / 209
Information for the Three-Parameter Model / 216
Metric Transformation, 3PL Model / 220
Handling Missing Responses / 220
Issues to Consider in Selecting among the 1PL, 2PL, and 3PL Models / 224
Summary / 226

7 • Rasch Models for Ordered Polytomous Data 237

Conceptual Development of the Partial Credit Model / 238

Conceptual Parameter Estimation of the PC Model / 243

*Example: Application of the PC Model to a Reasoning Ability Instrument,
MMLE, flexMIRT / 244*

*Example: Application of the PC Model to a Reasoning Ability Instrument,
MMLE, mirt / 256*

The Rating Scale Model / 267

Conceptual Parameter Estimation of the RS Model / 272

*Example: Application of the RS Model to an Attitudes Toward Condoms Scale,
JMLE, BIGSTEPS / 272*

*Example: Application of the PC Model to an Attitudes Toward Condoms Scale,
JMLE, mixRasch / 287*

How Large a Calibration Sample? / 292

Information for the PC and RS Models / 294

Metric Transformation, PC and RS Models / 296

Summary / 296

8 • Non-Rasch Models for Ordered Polytomous Data 313

The Generalized Partial Credit Model / 313

*Example: Application of the GPC Model to a Reasoning Ability Instrument,
MMLE, flexMIRT / 318*

*Example: Application of the GPC Model to a Reasoning Ability Instrument,
MMLE, mirt / 321*

Conceptual Development of the Graded Response Model / 324

How Large a Calibration Sample? / 333

Information for Graded Data / 334

Metric Transformation, GPC and GR Models / 336

*Example: Application of the GR Model to an Attitudes Toward Condoms Scale,
MMLE, flexMIRT / 337*

*Example: Application of the GR Model to an Attitudes Toward Condoms Scale,
MMLE, mirt / 340*

Conceptual Development of the Continuous Response Model / 343

Summary / 351

9 • Models for Nominal Polytomous Data 356

Conceptual Development of the Nominal Response Model / 357

Information for the NR Model / 365

Metric Transformation, NR Model / 366

Conceptual Development of the Multiple-Choice Model / 366

How Large a Calibration Sample? / 368

*Example: Application of the NR Model to a General Science Test, MMLE,
mirt / 370*

Summary / 383

10 • Models for Multidimensional Data	391
Conceptual Development of a Multidimensional IRT Model /	391
Multidimensional Item Location and Discrimination /	397
Item Vectors and Vector Graphs /	401
The Multidimensional Three-Parameter Logistic Model /	404
Assumptions of the MIRT Model /	404
Estimation of the M2PL Model /	405
Information for the M2PL Model /	406
Indeterminacy in MIRT /	408
Metric Transformation, M2PL Model /	410
<i>Example: Calibration of Interpersonal Engagement Instrument, M2PL Model,</i>	
<i>sirt.noharm /</i>	411
Obtaining Person Location Estimates /	421
<i>Example: Calibration of Interpersonal Engagement Instrument, M2PL Model,</i>	
<i>mirt /</i>	422
<i>Example: Calibration of Interpersonal Engagement Instrument, M2PL Model,</i>	
<i>flexMIRT /</i>	429
Summary /	431
 11 • Linking and Equating	 443
Equating Defined /	443
Equating: Data Collection Phase /	445
Equating: Transformation Phase /	446
<i>Example: Application of the Total Characteristic Function Equating Method,</i>	
<i>EQUATE /</i>	454
<i>Example: Application of the Total Characteristic Function Equating Method,</i>	
<i>SNSequate /</i>	463
<i>Example: Fixed-Item and Concurrent Calibration Equating /</i>	465
Summary /	471
 12 • Differential Item Functioning	 478
Differential Item Functioning and Item Bias /	479
Mantel–Haenszel Chi-Square /	483
The TSW Likelihood Ratio Test /	486
Logistic Regression /	487
<i>Example: DIF Analysis of Vocabulary Test, SAS CMH /</i>	491
<i>Example: DIF Analysis of Vocabulary Test, mantelhaen.test and</i>	
<i>difR /</i>	494
<i>Example: DIF Analysis of Vocabulary Test, SAS proc logistic /</i>	501
<i>Example: DIF Analysis of Vocabulary Test, glm and difR /</i>	508
Summary /	518

13 • Multilevel IRT Models

525

Multilevel IRT—Two Levels / 525

*Example: Estimating the Rasch Model from a Multilevel Perspective,**proc glimmix / 530**Example: Rasch Model Estimation, lme4 / 541*

Person-Level Predictors for Items / 545

*Example: Person-Level Predictors for Items—DIF Analysis,**proc glimmix / 547**Example: Person-Level Predictors for Items—DIF Analysis, lme4 / 551*

Person-Level Predictors for Respondents / 556

*Example: Person-Level Predictors for Respondents—Nutrition Literacy,**proc glimmix / 558**Example: Person-Level Predictors for Respondents, lme4 / 562*

Item-Level Predictors for Items / 567

*Example: Item-Level Predictors for Items—Nutrition Literacy,**proc glimmix / 569**Example: Item-Level Predictors for Items—Nutrition Literacy, lme4 / 571*

Multilevel IRT—Three Levels / 574

*Example: Three-Level Model Analysis—Nutrition Literacy,**proc glimmix / 579**Example: Three-Level Analysis of Nutrition Literacy Data, lme4 / 582*

Summary / 587

Appendices A–G can be accessed online at the book's companion website (www.guilford.com/deayala-materials), which also provides links to data, syntax, and output files in different software packages for the book's examples.

Appendix A. Maximum Likelihood Estimation of Person Locations

Estimating an Individual's Location: Empirical Maximum Likelihood Estimation

Estimating an Individual's Location: Newton's Method for MLE

R Function for MLE of θ with the Rasch Model

Revisiting Zero Variance Binary Response Patterns

Appendix B. Maximum Likelihood Estimation of Item Locations

R function for MLE of δ with the Rasch Model

Appendix C. The Normal Ogive Models

Conceptual Development of the Normal Ogive Model

The Relationship between IRT Statistics and Traditional Item Analysis Indices

Relationship of the Two-Parameter Normal Ogive and Logistic Models

Extending the Two-Parameter Normal Ogive Model to a Multidimensional Space

Appendix D. Computerized Adaptive Testing

A Brief History

Fixed-Branching Techniques

Variable-Branching Techniques

Advantages of Variable-Branching over Fixed-Branching Methods	
IRT-Based Variable-Branching Adaptive Testing Algorithm	
Appendix E. Linear Logistic Test Model (LLTM)	
Example of LLTM Calibration Using eRm	
Appendix F. Mixture Models	
Latent Class Analysis	
Mixture Rasch Model	
Example: Application of the Mixture Rasch Model to Writing Problem Data, CMLE, WINMIRA	
Example: Application of the Mixture Rasch Model to Writing Problem Data, CMLE, psychomix	
Appendix G. Miscellanea	
Using Principal Axis for Estimating Item Discrimination	
Infinite Item Discrimination Parameter Estimates	
Example: NOHARM Unidimensional Calibration	
An Approximate Chi-Square Statistic for NOHARM	
Relative Efficiency, Monotonicity, and Information	
FORTTRAN Formats	
Odds, Odds Ratios, and Logits	
The Person Response Function	
Linking: A Temperature Analogy Example	
Should DIF Analyses Be Based on Latent Classes?	
The Separation and Reliability Indices	
Dependency in Traditional Item Statistics and Observed Scores	
Conditional Independence Using Q_3	
Standalone NOHARM Calibration of Interpersonal Engagement Instrument, M2PL Model	
CFI, GFI, M_2 , RMSEA, TLI, and SRMR	
An Introduction to Kernel Equating	
Correspondence between the Rasch Model and a Loglinear Model	
R Introduction	

References	597
Author Index	625
Subject Index	631
About the Author	643