

Contents

Preface	xiii
Acknowledgment	xv
I Theory	1
1 Introduction	3
1.1 The Role of Psychological and Educational Tests	3
1.2 The Rasch Model and Item Response Theory	4
1.3 Where You Will Find What in This Book	6
2 The Rasch Model	9
2.1 The Data Matrix	10
2.2 The Item Response Function	12
2.2.1 Ability and Difficulty	12
2.2.2 Discrimination	14
2.2.3 The Logistic Function	16
2.3 Alternative Representations	18
2.3.1 Probability of an Incorrect Response	18
2.3.2 Probability of an Arbitrary Response	19
2.3.3 Alternative Representation of the Logistic Function	20
2.3.4 Multiplicative Form	20
2.4 Properties and Assumptions	21
2.4.1 Sufficient Statistics	21
2.4.2 Local Stochastic Independence	23
2.4.2.1 Items	24
2.4.2.2 Persons	26
2.4.3 Specific Objectivity	28
2.4.4 Unidimensionality	32
2.4.5 Measurement Scale	32
2.5 Exercises	34
3 Parameter Estimation	37
3.1 Joint Maximum Likelihood Estimation	38
3.2 Conditional Maximum Likelihood Estimation	39
3.3 Marginal Maximum Likelihood Estimation	43

3.4	Bayesian Estimation	45
3.5	Person Parameter Estimation	48
3.6	Item and Test Information	50
3.7	Sample Size Requirements	53
3.8	Exercises	54
4	Test Evaluation	57
4.1	Graphical Assessment	59
4.1.1	Person Item Map	59
4.1.2	Empirical ICCs	60
4.1.3	Graphical Test	62
4.2	Tests for Item and Person Invariance	64
4.2.1	Andersen's Likelihood Ratio Test	65
4.2.2	Martin-Löf Test and Other Approaches for Detecting Multidimensionality	67
4.2.3	Wald Test	68
4.2.4	Anchoring	69
4.2.5	Other Approaches for Detecting DIF	71
4.2.6	How to Proceed with Problematic Items	71
4.3	Goodness-of-Fit Tests and Statistics	72
4.3.1	χ^2 and G^2 Goodness-of-Fit Tests	72
4.3.2	M_2 , RMSEA, and SRMSR	74
4.3.3	Infit and Outfit Statistics	75
4.3.4	Further Fit Statistics for Items	79
4.3.5	Fit Statistics for Item Pairs	81
4.3.6	Fit Statistics for Persons	82
4.3.7	Nonparametric Goodness-of-Fit Tests	83
4.3.8	Posterior Predictive Checks	84
4.4	Separation Indices	85
4.4.1	Item Separation Index	85
4.4.2	Person Separation Index	86
4.5	Evaluation Through Model Comparisons	87
4.5.1	Models with Additional Item Parameters	87
4.5.1.1	Two-Parameter Model	87
4.5.1.2	Three-Parameter Model	89
4.5.1.3	Four-Parameter Model	91
4.5.1.4	Sample Size Requirements	91
4.5.2	Likelihood Ratio Tests	92
4.5.3	Information Criteria	93
4.6	Exercises	94

II Applications	97
5 Basic R Usage	99
5.1 Installation of R and Add-On Packages	99
5.2 Code Editors and RStudio	101
5.3 Loading and Importing Data	102
5.4 Getting Information About Persons and Variables	103
5.5 Addressing Elements in Lists	108
5.6 Exercises	110
6 R Package eRm	111
6.1 Item Parameter Estimation	112
6.2 Test Evaluation	119
6.2.1 Person Item Map	120
6.2.2 Empirical ICCs	121
6.2.3 Andersen's Likelihood Ratio Test and Graphical Test	121
6.2.4 Wald Test	127
6.2.5 Anchoring	128
6.2.6 Removing Problematic Items	133
6.2.7 Martin-Löf Test	134
6.2.8 Item and Person Fit	135
6.3 Plots of ICCs, Item and Test Information	137
6.4 Person Parameter Estimation	139
6.5 Test Evaluation in Small Data Sets	142
6.6 Exercises	145
7 R Package mirt	147
7.1 Model Selection	148
7.2 Item Parameter Estimates	150
7.2.1 Illustration via Expected ICCs	150
7.2.2 Displaying the Estimates	152
7.3 Evaluating Goodness-of-Fit	156
7.4 Ability Estimation	160
7.5 Exercises	163
8 R Package TAM	165
8.1 Item Parameter Estimation	166
8.2 Evaluating Goodness-of-Fit	168
8.3 Person Parameter Estimation	170
8.4 Exercises	172
9 R Interface to Stan	173
9.1 Stan Models	174
9.1.1 The data Block	174
9.1.2 The parameters Block	176

9.1.3	The transformed parameters Block	176
9.1.4	The model Block	177
9.2	Sampling the Posterior Using RStan	178
9.3	Evaluating Goodness-of-Fit	183
9.4	Exercises	189
Summary of R Commands for eRm, mirt, and TAM		191
III Beyond the Rasch Model		195
10	Extensions to the Rasch Model	197
10.1	The Linear-Logistic Test Model	197
10.2	Modeling Differences Between People	199
10.2.1	The Mixture Rasch Model	199
10.2.2	Model-Based Recursive Partitioning	200
10.2.3	Explanatory IRT – The Rasch Model as a Mixed Model	201
10.3	Multidimensional IRT Models	203
10.4	Exercises	203
11	Models for Polytomous Responses	205
11.1	The Partial Credit Model	206
11.1.1	CCCs and Threshold Parameters	209
11.1.2	Alternative Parameterizations	210
11.1.3	Disordered Thresholds	214
11.2	The Rating Scale Model	215
11.3	The Generalized Partial Credit and the Nominal Response Model	217
11.4	The Graded Response Model	217
11.5	The Sequential Model	218
11.6	Sample Size Requirements	219
11.7	Exercises	219
11.8	Derivations for the Partial Credit Model	220
12	Outlook on Special Applications	223
12.1	Computerized Adaptive Testing	223
12.2	Test Linking and Equating	224
12.3	Longitudinal IRT Models	225
12.4	Exercises	226
Appendices		227
A	Useful Mathematical Formulas	229
A.1	Sums and Products	229
A.2	Exponentials	230
A.3	Logarithms	230

A.4	Differentiation Rules	231
A.4.1	Rules	231
A.4.2	Examples	232
B	Statistical Background	233
B.1	Statistical Estimation	233
B.1.1	The Binomial Distribution	233
B.1.2	Maximum Likelihood Estimation	235
B.1.3	Likelihood for Multiple Observations	238
B.1.4	Bayesian Inference	241
B.1.4.1	Bayes' Rule by Example	241
B.1.4.2	Coin Flipping with a Uniform Prior	243
B.1.4.3	Informative Priors and Beta-Binomial Model	246
B.2	Statistical Testing	248
B.2.1	Tests Based on the χ^2 Distribution	249
B.2.1.1	χ^2 Test for Independence	249
B.2.1.2	Goodness-of-Fit and Other χ^2 Tests	252
B.2.2	Tests Based on the Normal Distribution	253
C	Answers to the End of Chapter Questions	255
C.2	Answers for Chapter 2	255
C.3	Answers for Chapter 3	256
C.4	Answers for Chapter 4	259
C.5	Answers for Chapter 5	262
C.6	Answers for Chapter 6	265
C.7	Answers for Chapter 7	274
C.8	Answers for Chapter 8	281
C.9	Answers for Chapter 9	283
C.10	Answers for Chapter 10	286
C.11	Answers for Chapter 11	287
C.12	Answers for Chapter 12	288
	References	289
	Author Index	301
	Index	305