

# Contents

<b>1</b>	<b>Theories of Data Analysis and Statistical Inference</b>	<b>1</b>
1.1	Introduction	1
1.2	Example	2
1.3	Statistical models	2
1.4	The likelihood function	5
1.5	Theories	6
1.5.1	Likelihood-based repeated sampling theory	6
1.5.2	Bayes theory	6
1.5.3	“Model assisted” survey sampling theory	8
1.6	Weighting	13
1.6.1	Stratified random sampling	13
1.6.2	Design-based analysis	14
1.6.3	Model-based analysis	15
1.6.4	Weighted likelihoods	16
1.7	Missing data and non-response	18
1.7.1	Weighting adjustments for nonresponse	19
1.7.2	Incomplete data in regression	20
1.7.3	Multiple imputation	20
<b>2</b>	<b>The Current Design and Analysis</b>	<b>23</b>
2.1	NCES and NAEP	23
2.2	Design	24
2.2.1	PSUs	24
2.2.2	Schools	25
2.2.3	Students	25
2.2.4	Test items	25
2.2.5	Important design issues	26
2.3	NAEP state sample design 2002+	26
2.4	Weighting	26
2.4.1	Design effect corrections	27
2.5	Analysis	28

2.5.1	Item models . . . . .	28
2.5.2	Multidimensional ability . . . . .	30
2.5.3	Inference and the likelihood function . . . . .	34
2.5.4	The ability regression model . . . . .	35
2.5.5	Current model parameter estimation . . . . .	36
2.5.6	Plausible value imputation . . . . .	36
<b>3</b>	<b>Psychometric and Survey Models . . . . .</b>	<b>39</b>
3.1	The Rasch model . . . . .	39
3.2	The 2PL and MIMIC models . . . . .	40
3.3	Three-parameter models . . . . .	42
3.4	Partial credit model . . . . .	44
3.5	The HYBRID model . . . . .	44
3.6	Extensions of the guessing model . . . . .	45
3.6.1	A four-parameter guessing model . . . . .	45
3.6.2	The “2-guess” model . . . . .	46
3.6.3	The “2-mix” model – a five-parameter general mixture of logits model . . . . .	46
3.7	Modeling the component membership probability . . . . .	48
3.8	Multidimensional ability . . . . .	48
3.9	Clustering and variance component models . . . . .	50
3.9.1	Three-level models . . . . .	51
3.9.2	Four-level models . . . . .	52
3.10	Summary of the full model for NAEP analysis . . . . .	53
<b>4</b>	<b>Technical Reports – Data Analyses and Simulation Studies . . . . .</b>	<b>55</b>
4.1	Research reports . . . . .	55
<b>5</b>	<b>1986 NAEP Math Survey . . . . .</b>	<b>63</b>
5.1	Data and model specification – subscale . . . . .	63
5.2	Model aspects . . . . .	65
5.2.1	Maximised log-likelihoods . . . . .	65
5.2.2	Two- and three-level models . . . . .	65
5.2.3	Four-level models . . . . .	66
5.2.4	The 3PL model . . . . .	66
5.3	Reporting group differences . . . . .	67
5.3.1	Comparison with NAEP subscale estimates . . . . .	69
5.4	Mixture models . . . . .	71
5.4.1	2-guess model . . . . .	72
5.4.2	2-guess-prob model . . . . .	72
5.4.3	Two-dimensional model . . . . .	73
5.4.4	2-mix model . . . . .	74
5.4.5	2-mix-regressions model . . . . .	74
5.4.6	2-mix-prob model . . . . .	74
5.4.7	Conclusions from the 30-item analysis . . . . .	75

<b>6</b>	<b>Analysis of All 1986 Math Items</b>	77
6.1	The full math test	77
6.2	2PL models	78
6.3	Results	79
6.3.1	Mixed 2PL models	80
6.3.2	Three-component membership models	81
6.3.3	Multidimensional ability model	82
6.4	MIMIC models	83
6.5	Results	83
6.5.1	Two-parameter MIMIC model	83
6.5.2	Mixed MIMIC models	84
6.6	Comparison with published NAEP results	85
6.7	Discussion	86
<b>7</b>	<b>2005 NAEP Math Survey – Texas</b>	87
7.1	Population, sample, and test	87
7.2	Variable names and codes	88
7.3	Models fitted	88
7.4	Results – limited teacher data	90
7.4.1	Three-parameter interpretation	90
7.4.2	Mixture models	91
7.5	Boundary values in logistic regression	93
7.6	Results – extensive teacher data	93
7.6.1	Mixture models	95
7.7	Comparison with official NCES analysis	96
7.8	Conclusion	98
<b>8</b>	<b>2005 NAEP Math Survey – California</b>	99
8.1	Population, sample, and test	99
8.2	Models	100
8.3	Results	100
8.3.1	Limited teacher data	100
8.3.2	3PL interpretation	101
8.4	Mixture models	102
8.5	Extensive teacher data	103
8.5.1	3PL interpretation	104
8.6	Mixture models	105
8.7	Comparison with official NCES analysis	107
8.8	Conclusion	109
<b>9</b>	<b>Conclusions</b>	111
9.1	The nature and structure of models	111
9.2	Our modeling results	112
9.2.1	Comparisons with published NAEP tables	112
9.2.2	Main effects and interactions	112

9.2.3	Mixtures and latent subpopulations . . . . .	113
9.3	Current analysis . . . . .	116
9.3.1	Dependence of design on analysis . . . . .	116
9.3.2	Multilevel modeling . . . . .	117
9.3.3	The limitations of NAEP data for large-scale modeling . . . . .	117
9.4	The reporting of NAEP data . . . . .	118
9.5	The future analysis and use of NAEP data . . . . .	119
9.6	Resolution of the model-comparison difficulties . . . . .	120
9.7	Resolution of the problems with incomplete data . . . . .	120
<b>A</b>	<b>1986 Survey Results, 30 Item Subscale . . . . .</b>	<b>121</b>
<b>B</b>	<b>1986 Survey Results, Full 79 Items . . . . .</b>	<b>133</b>
<b>C</b>	<b>Model Parameter Estimates and SEs, 2005 Texas Survey . . . . .</b>	<b>141</b>
C.1	Parameter estimates and SEs – limited teacher data . . . . .	146
C.2	Parameter estimates and SEs – extensive teacher data . . . . .	148
<b>D</b>	<b>Model Parameter Estimates and SEs, 2005 California survey . . . . .</b>	<b>149</b>
D.1	Parameter estimates for MIMIC models – limited teacher data . . . . .	149
D.2	Parameter estimates for MIMIC models – extensive teacher data . . . . .	153
	<b>References . . . . .</b>	<b>157</b>
	<b>Author Index . . . . .</b>	<b>161</b>

Statistical Modeling of the National Assessment of  
Educational Progress

Aitkin, M.; Aitkin, I.

2011, XII, 164 p., Hardcover

ISBN: 978-1-4419-9936-8