Universität der Bundeswehr München

Eine Methode zur Parameterbestimmung im Rasch-Modell bei fehlenden Werten

Jörg-Henrik Heine, Christian Tarnai & Florian Hartmann
Universität der Bundeswehr München,
Fakultät für Pädagogik,
Werner-Heisenberg-Weg 39, 85577 Neubiberg
joerg-henrik.heine@unibw.de

10. Tagung der Fachgruppe Methoden & Evaluation der DGPs
vom 21.09.2011 bis zum 23.09.2011 in Bamberg

Einführung

- Parameterschätzung bei Raschskalierung psychometrischer Skalen
  - conditional Maximum Likelihood (cML)
    - Separate Schätzung von Personen- und Item Parametern
    - ermöglicht „biasfreie“ (verteilungsunabhängige) Parameterschätzung
  - joint Maximum Likelihood (jML)
    - Gleichzeitige Schätzung von Personen- und Item Parametern
    - Item Parameterschätzung nicht absolut „biasfrei“ insbesondere bei kurzen Skalen (Willse, 2010)
- Grundprinzip und Problematik bei beidem:
  - Parameter werden iterativ **geschätzt**
  - Bei systematisch fehlenden Werten mögliche verzerrte Parameterschätzung

Alternative zur Parameterschätzung

- Die **explizite Berechnung** der Item Parameter nach der Methode des „paarweisen Item Vergleichs“
- Prinzip:
  - Zählen der der „lösungs-“Häufigkeiten von Item *j* unter Voraussetzung das Item *i* nicht gelöst wurde ($i \neq j$)
  - Bildung der positiv reziproken Häufigkeits-Matrix
  - Potenzieren der positiv reziproken Häufigkeits-Matrix mit 3
  - Logarithmieren der potenzierten Matrix und bilden der Zeilenrandsummen
  - Einsetzen des Item Parameter-Vektors und rawscore-Vektors in die Rasch-Modellgleichung zur ML-Schätzung der Personen Parameter
- Vorteil:
  - Separate Berechnung/Schätzung von Item- und Personen Parametern
  - Direkt anwendbar auf Datenmatrizen mit fehlenden Werten

Vorgehen / Methode

- Raschskalierung einer 7 Item Skala „*Einstellung zu Schwangerschaftsabbruch*“ (ALLBUS)
  WinMira (Davier, 2001) → „list wise deletion” zur Erzeugung einer „vollständigen“ Datenmatrix
  - Getestete Modellannahme: PCM und mixed-PCM (Masters, 1982)Modellvergleich anhand BIC und CAIC
- Bestimmung der Item- und Personenparameter anhand der vollständigen Datenmatrix N = 2068 (WinMira)
- Bestimmung der Item- und Personenparameter anhand der gesamten, Datenmatrix N = 2739 (pairwise comparison)
- Simulation unterschiedlicher missing Anteile in Abhängigkeit einer weiteren Variable (MAR) und Bestimmung der Item und Personen Parameter anhand der unvollständigen Datenmatrix (pairwise comparison)

Theorie: Bestimmung Item Parameter σ

pairwise Algorithmus für dichotome Daten nach Choppin (1968, 1986)
vgl. auch Rasch (1980); Zwinderman (1995); Garner & Engelhard (2009)

	Item1	*Item2*	*Item3*	*Item4*	. . .
Person 1	1	1	1	0	
Person 2	0	1	0	1	
Person 3	1	0	0	1	
Person 4	1	0	0	0	
⋮					

Paarweiser Vergleich

n_{ij}

Items	***1***	***2***	***3***	***4***
1	0	2	2	2
2	1	0	1	1
3	0	0	0	1
4	1	1	2	0

n_{ij}^2

Items	***1***	***2***	***3***	***4***
1	4	2	6	4
2	1	3	4	3
3	1	1	2	0
4	1	2	3	5

n_{ij}^3

Items	***1***	***2***	***3***	***4***
1	6	12	18	16
2	6	5	11	9
3	1	2	3	5
4	7	7	14	7

pos. rez. Matrix

Items	***1***	***2***	***3***	***4***
1	1	6/12	1/18	7/16
2	12/6	1	2/11	7/9
3	18/1	11/2	1	14/5
4	16/7	9/7	5/14	1

Log

Items	***1***	***2***	***3***	***4***
1	0	-0,693	-2,891	-0,828
2	0,693	0	-1,709	-0,252
3	2,89	1,705	0	1,03
4	0,826	0,251	-1,03	0

Zeilen-randsummen

Items	***1***	***2***	***3***	***4***	
1	0	-0,693	-2,891	-0,828	**-4,412**
2	0,693	0	-1,709	-0,252	**-1,268**
3	2,89	1,705	0	1,03	**5,625**
4	0,826	0,251	-1,03	0	**0,047**

explizit berechnete Item Parameter σ

Theorie: Bestimmung der Personen Parameter θ (ML)

Modellgleichung (Raschmodell)

$$p(x_{vi}) = \frac{\exp(x_{vi}(\theta_v - \sigma_i))}{1+\exp(\theta_v - \sigma_i)}, x \in \{0,1\} \quad (1)$$

Wahrscheinlichkeit der Daten unter Modell Annahme

$$L = p(\underline{\underline{x}}) = \prod_{v=1}^{N}\prod_{i=1}^{k} p(x_{vi}), x \in \{0,1\} \quad (2)$$

Einsetzen von (1) in (2)

$$L = \prod_{v=1}^{N}\prod_{i=1}^{k} \frac{\exp(x_{vi}(\theta_v - \sigma_i))}{1+\exp(\theta_v - \sigma_i)}, x \in \{0,1\} \quad (3)$$

Umformen und ausmultiplizieren des Zählers von (3)

$$L = \frac{\exp\left(\sum_{v=1}^{N}\sum_{i=1}^{k} x_{vi}\theta_v - \sum_{v=1}^{N}\sum_{i=1}^{k} x_{vi}\sigma_i\right)}{\prod_{v=1}^{N}\prod_{i=1}^{k} 1+\exp(\theta_v - \sigma_i)}, x \in \{0,1\} \quad (4)$$

Ersetzen der Summen über *k* in (4)

$$L = \frac{\exp\left(\sum_{v=1}^{N} r_v\theta_v - \sum_{v=1}^{N} n_i\sigma_i\right)}{\prod_{v=1}^{N}\prod_{i=1}^{k} 1+\exp(\theta_v - \sigma_i)}, x \in \{0,1\} \quad (5)$$

Logarithmieren von (5)

$$\log L = \sum_{v=1}^{N} r_v\theta_v - \sum_{i=1}^{k} n_i\sigma_i - \sum_{v=1}^{N}\sum_{i=1}^{k} \log(1+\exp(\theta_v - \sigma_i)), x \in \{0,1\} \quad (6)$$

Erste partielle Ableitung nach *θ* ...

...und umformen von (6)

$$\hat{\theta}_v = \log(r_v) - \log \sum_{i=1}^{k} \frac{\exp(-\sigma_i)}{1+\exp(\theta_v - \sigma_i)}$$

Iterative Schätzung von **θ** mit explizit berechnetem ***σ*** und bekanntem $\boldsymbol{r_v}$

Daten / Stichprobe

- **All**gemeine **B**evölkerungs **U**mfrage der **S**ozialwissenchaft (**ALLBUS**)
- Erhebungsjahr 1982
- Insgesamt: N = 2739
- Alter: 18 - 89 Jahre, MW = 46.06, MD = 44, SD = 17.05
- n_1 = 1230 männlich und n_2 = 1509 weiblich

Variablen

7 dichotome Items „*Einstellung zu Schwangerschaftsabbruch*“

Bitte sagen Sie mir, ob es Ihrer Meinung nach einer Frau gesetzlich möglich sein sollte oder nicht, einen Schwangerschaftsabbruch vornehmen zu lassen ...

		INT.: Aussagen vorlesen und Antwortziffer einkreisen	ja, sollte möglich sein	nein, sollte nicht möglich sein	weiß nicht
Item 1	A	...wenn das Baby mit hoher Wahrscheinlichkeit eine ernsthafte Schädigung haben wird?	1	2	8
Item 2	B	...wenn die Frau verheiratet ist und keine Kinder mehr haben möchte?	1	2	8
Item 3	C	...wenn die Gesundheit der Frau durch die Schwangerschaft ernsthaft gefährdet ist?	1	2	8
Item 4	D	...wenn die Familie nur über ein geringes Einkommen verfügt und sich keine Kinder mehr leisten kann?	1	2	8
Item 5	E	...wenn die Schwangerschaft Folge einer Vergewaltigung ist?	1	2	8
Item 6	F	...wenn die Frau unverheiratet ist und den Vater des Kindes nicht heiraten möchte?	1	2	8
Item 7	G	...wenn die Frau es so will, unabhängig davon, welchen Grund sie dafür hat?	1	2	8

Rekodierung → 0 1 NA

Vollständige Daten für die Skala „*Einstellung zu Schwangerschaftsabbruch*“

N = 2068

Alter

Range: 18 - 88 Jahre,

MW = 45.14,

MD = 43,

SD = 16.87

n_1 = 941 männlich

n_2 = 1127 weiblich

Variablen

Häufigkeit des Kirchgangs ?

11	Wie oft gehen Sie im allgemeinen zur Kirche: *INT.: Antwortvorgaben vorlesen*	mehr als einmal in der Woche . . 1 einmal in der Woche . . . . . 2 ein- bis dreimal im Monat . . . 3 mehrmals im Jahr . . . . . 4 seltener . . . . . 5 nie 6 9	17 13

Rekodierung

1 = 1 „mehr als einmal die Woche“
2 = 2
3 = 3
4 = 4
5 = 5
6 = 6 „nie“
7 = → „nie, aus religiösen Gründen (nicht Kath. o. Ev.)“

N = 2739	Häufigkeit des Kirchgangs ?						
Antwortkategorie	1	2	3	4	5	6	7
Prozent Kategorie	3,36	13,58	11,24	21,58	27,35	14,82	8,07
Häufigkeit Kategorie	92	372	308	591	749	406	221

N = 2068	Häufigkeit des Kirchgangs ?						
Antwortkategorie	1	2	3	4	5	6	7
Prozent Kategorie	3,38	12,52	10,54	21,47	27,37	15,81	8,9
Häufigkeit Kategorie	70	259	218	444	566	327	184

Empirie - Modellvergleich

Ein-Klassenlösung

Goodness of fit statistics:

		estimated model	saturated model
Log-Likelihood	:	-5084.84	-4977.51
Number of parameters	:	13	127
geom. mean likelihood	:	0.70380085	0.70903825

Information Criteria:

AIC-Index	:	10195.67	10209.02
BIC-Index	:	10268.92	10924.58
CAIC-Index	:	10281.92	11051.58

class	size	prob.		
1	0.706	0.959	0.959	0.041
2	0.294	0.708	0.292	0.708

Goodness of fit statistics:

		estimated model	saturated model
Log-Likelihood	:	-5064.54	-4977.51
Number of parameters	:	25	127
geom. mean likelihood	:	0.70478820	0.70903825

Information Criteria:

AIC-Index	:	10179.09	10209.02
BIC-Index	:	10319.95	10924.58
CAIC-Index	:	10344.95	11051.58

Zwei-Klassenlösung

Ausfallmechanismus

	Alter	Geschlecht	Kirchgang
missing auf Item 1	.05	.00	-.13
missing auf Item 2	.04	.02	.02
missing auf Item 3	.08	.06	-.08
missing auf Item 4	.06	.02	-.03
missing auf Item 5	.05	.02	-.16
missing auf Item 6	.03	.00	.05
missing auf Item 7	.03	-.01	.03

Anmerkung. Kirchgang ist negativ gepolt

→ Je höher die Kirchgangshäufigkeit desto eher fehlen Daten auf Items mit Inhalten zu physischer Versehrtheit als Begründung für den Schwangerschaftsabbruch

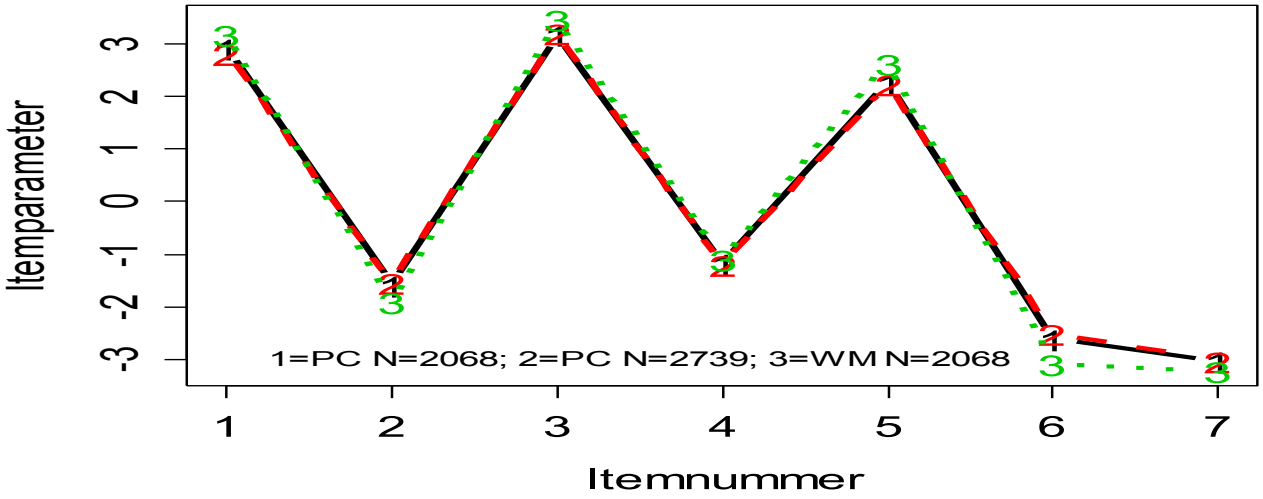
Kovariate zu Antwortendenzen

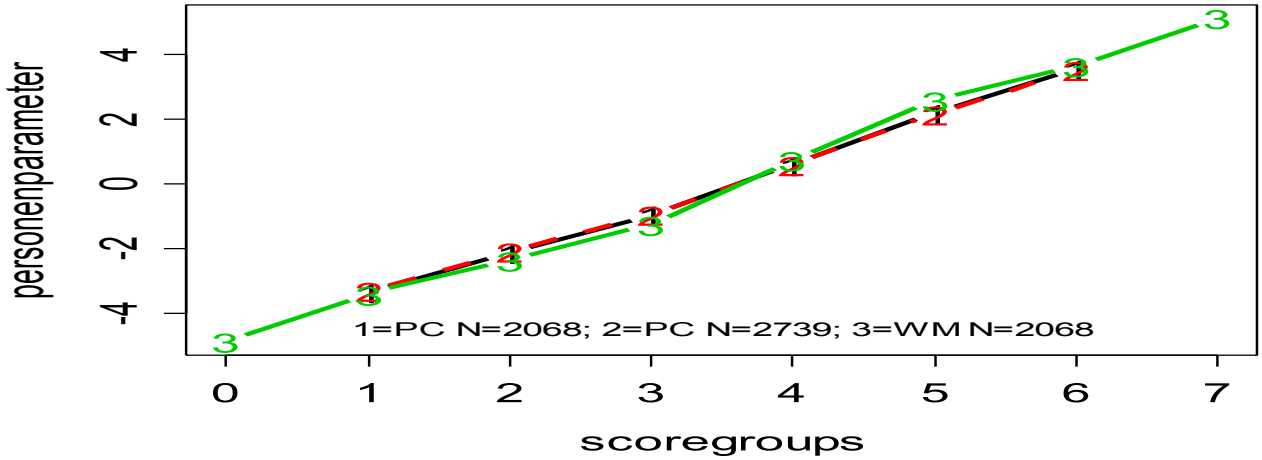
	Alter	Geschlecht	Kirchgang
Zustimmung Item 1	-0,08	-0,03	0,25
Zustimmung Item 2	-0,17	-0,02	0,38
Zustimmung Item 3	-0,11	-0,05	0,22
Zustimmung Item 4	-0,15	-0,01	0,37
Zustimmung Item 5	-0,09	-0,03	0,25
Zustimmung Item 6	-0,17	-0,04	0,34
Zustimmung Item 7	-0,18	-0,03	0,34

Anmerkung. Kirchgang ist negativ gepolt;
Itemzustimmung = 0, Itemablehnung = 1

→ Je niedriger die Kirchgangshäufigkeit desto eher wird den Items der Skala „*Einstellung zum Schwangerschaftsabbruch*“ zugestimmt

Empirie – Parameterbestimmung





1 = pairwise comp.

N = 2068

2 = pairwise comp.

N = 2739

3 = WinMira

(cML-Schätzung)

N = 2068

Fazit Empirie

- Rasch-Skalierung:
  - **Ein**-Klassenlösung für die Skala „*Einstellung zu Schwangerschaftsabbruch*“
- Vergleich der Parameterbestimmung
  - Vollständige Datenmatrix (list wise) und gesamte Datenmatrix (missings) als Basis ergeben vergleichbare Item- und Personenparameter
  - Beide Methoden (cML vs. pairwise) resultieren in vergleichbaren Item- und Personenparametern
- Ausfallmechanismus (Skala Schangerschaftsabbruch)
  - Die Datenausfallwahrscheinlichkeit lässt sich in Abhängigkeit der Variable Kirchgangshäufigkeit modellieren

Simulation

- Ausgehend von der vollständigen Datenmatrix (*N* = 2068) Simulation von 12 Datensätzen
- Zunehmender Anteil (von 5% bis 60%) an fehlenden Werten auf den 7 Items „*Einstellung zum Schwangerschaftsabbruch*“
- Modellierung der Datenausfallwahrscheinlichkeit in Abhängigkeit der „trait“ Ausprägung auf der Variable **Kirchgangshäufigkeit**
- **Für jeden der 12 Datensätze ...**
  - Berechnung der Item- und Personen Parameter mit **WinMira** (ein Klassen PCM)
    - Basis: Jeweils die „vollständige“ Datenmatrize mit abnehmendem *N* (list wise deletion)
  - Berechnung der Item- und Personen Parameter mit **pairwise comparison** (ein Klassen PCM)
    - Basis: Jeweils die gesamte Datenmatrize (*N* = 2068) mit zunehmendem Anteil missing

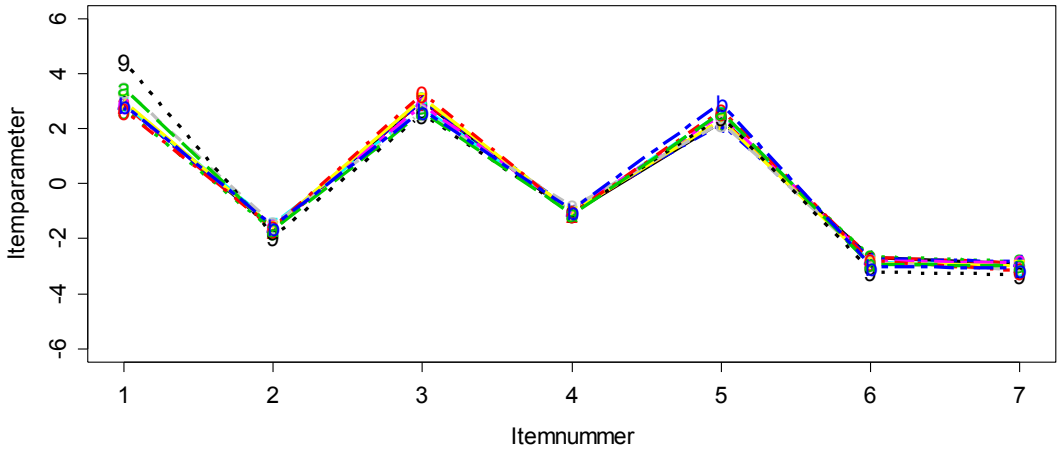
Datensätze Simulation

Simulierter Datensatz Nr.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
missing Anteil in Prozent	5	10	15	20	25	30	35	40	45	50	55	60
N (listwise)	1463	1062	730	539	369	277	175	136	103	66	39	38
N gesamt	2068	2068	2068	2068	2068	2068	2068	2068	2068	2068	2068	2068
Häufigkeit rawscore group 0	404	320	243	188	126	112	69	65	36	28	16	20
Häufigkeit rawscore group 1	117	75	65	41	37	22	13	14	14	3	5	6
Häufigkeit rawscore group 2	180	141	96	62	50	35	26	20	13	6	5	4
Häufigkeit rawscore group 3	221	153	112	83	45	30	18	9	14	11	3	3
Häufigkeit rawscore group 4	380	273	162	128	86	59	43	21	18	12	7	3
Häufigkeit rawscore group 5	85	48	31	18	13	10	3	3	2	3	2	1
Häufigkeit rawscore group 6	28	18	9	10	3	1	1	-	1	-	-	-
Häufigkeit rawscore group 7	48	34	12	9	9	8	2	4	5	3	1	1
Korr. Kirchgang und missing Item 1	-.14	-.13	-.18	-.20	-.19	-.24	-.24	-.27	-.3	-.33	-.32	-.30
Korr. Kirchgang und missing Item 2	-.10	-.17	-.19	-.18	-.18	-.23	-.27	-.26	-.29	-.32	-.31	-.33
Korr. Kirchgang und missing Item 3	-.09	-.14	-.17	-.19	-.22	-.24	-.26	-.29	-.26	-.29	-.31	-.32
Korr. Kirchgang und missing Item 4	-.08	-.17	-.17	-.21	-.25	-.23	-.23	-.28	-.27	-.34	-.3	-.33
Korr. Kirchgang und missing Item 5	-.09	-.15	-.17	-.18	-.22	-.25	-.23	-.26	-.32	-.32	-.31	-.34
Korr. Kirchgang und missing Item 6	-.11	-.15	-.15	-.17	-.2	-.28	-.25	-.31	-.28	-.29	-.33	-.33
Korr. Kirchgang und missing Item 7	-.06	-.15	-.14	-.20	-.22	-.22	-.24	-.27	-.29	-.29	-.28	-.35

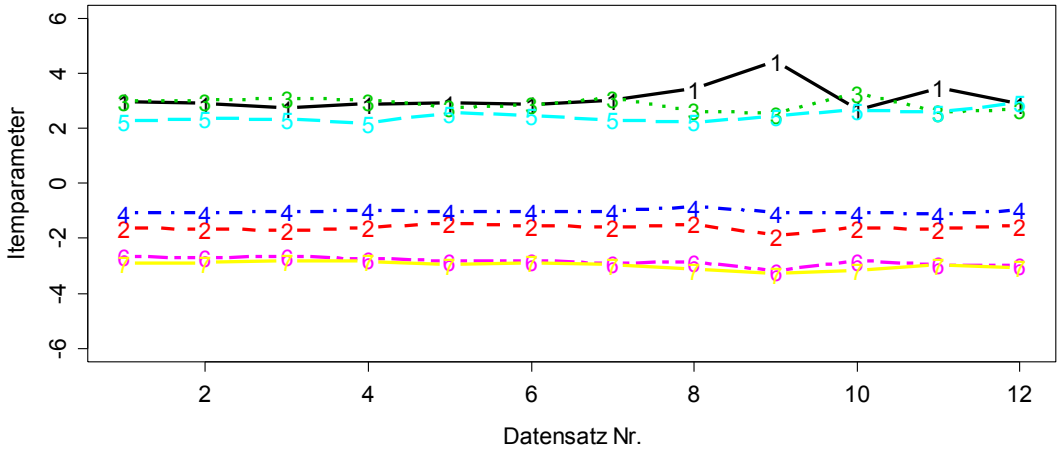
Anmerkung. Kirchgang ist negativ gepolt

Parameterbestimmung Simulation pairwise comparison

1. Item Parameter Basis: mit 3 potenzierte nij Matrix



Linien = Datensätze 1-12

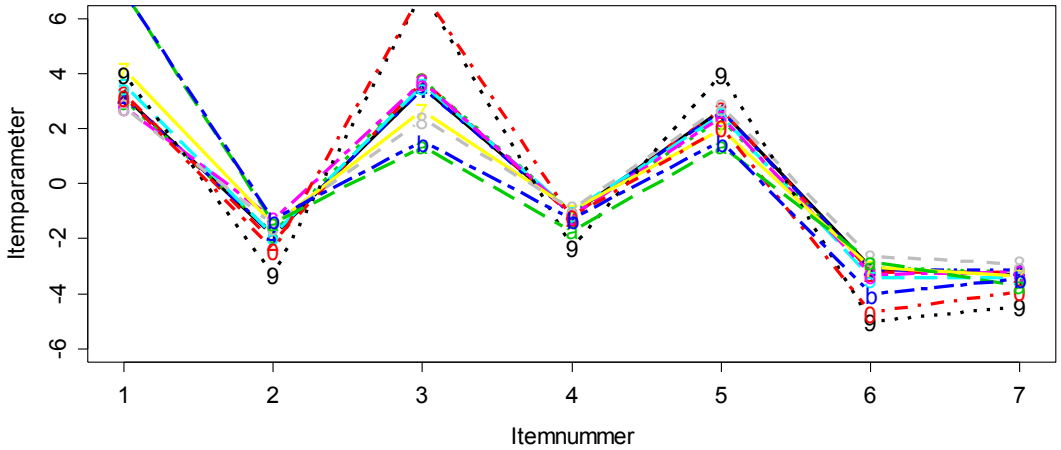


Linien = items 1-7

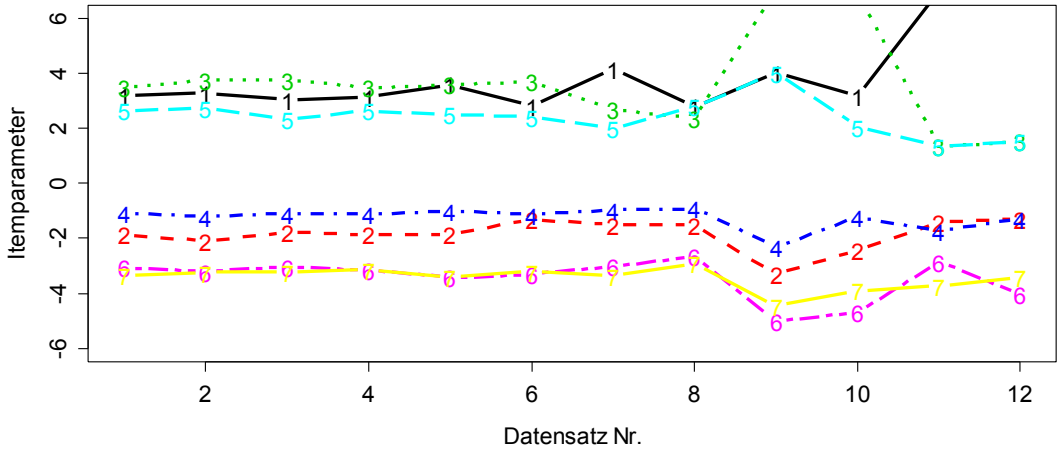
5	10	15	20	25	30	35	40	45	50	55	60	missing Anteil in Prozent

Parameterbestimmung Simulation WinMira (cML)

1. Item Parameter



Linien = Datensätze 1-12

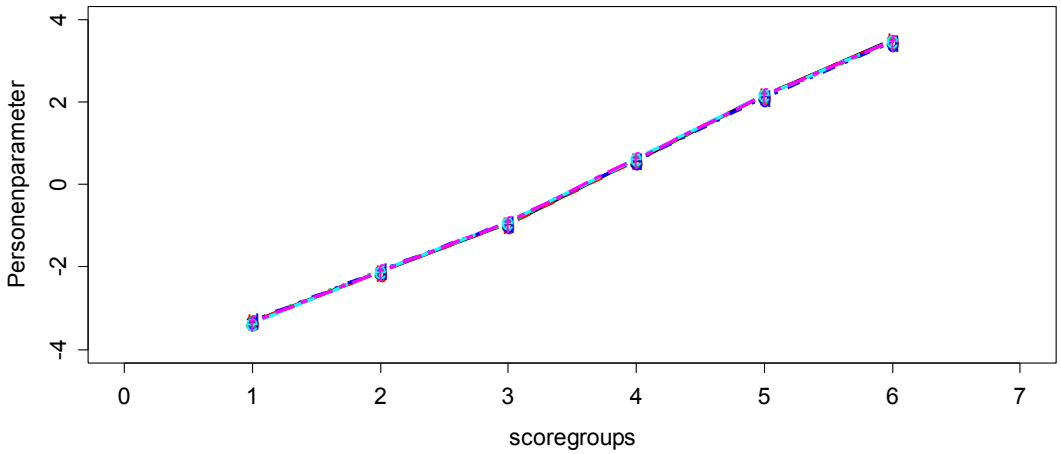


Linien = items 1-7

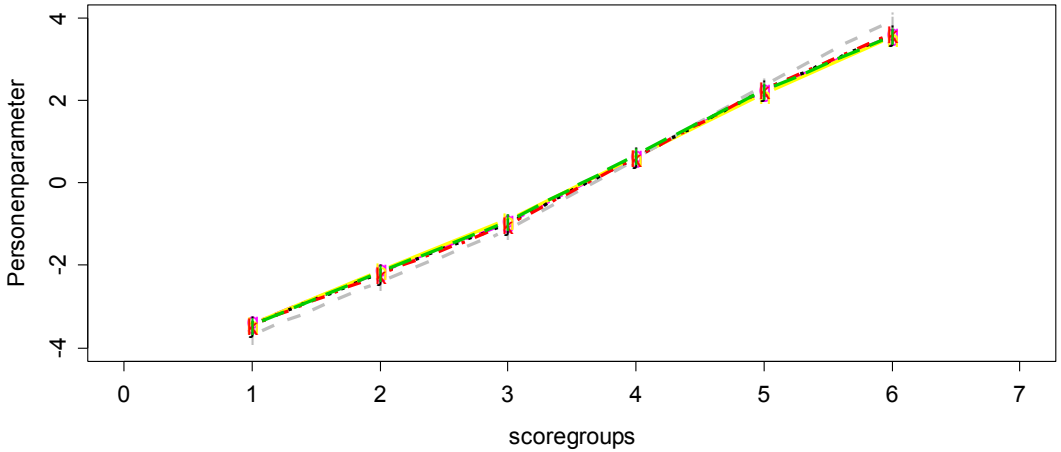
5	10	15	20	25	30	35	40	45	50	55	60	missing Anteil in Prozent

Parameterbestimmung Simulation pairwise comparison

2. Personen Parameter Basis: mit 3 potenzierte nij Matrix



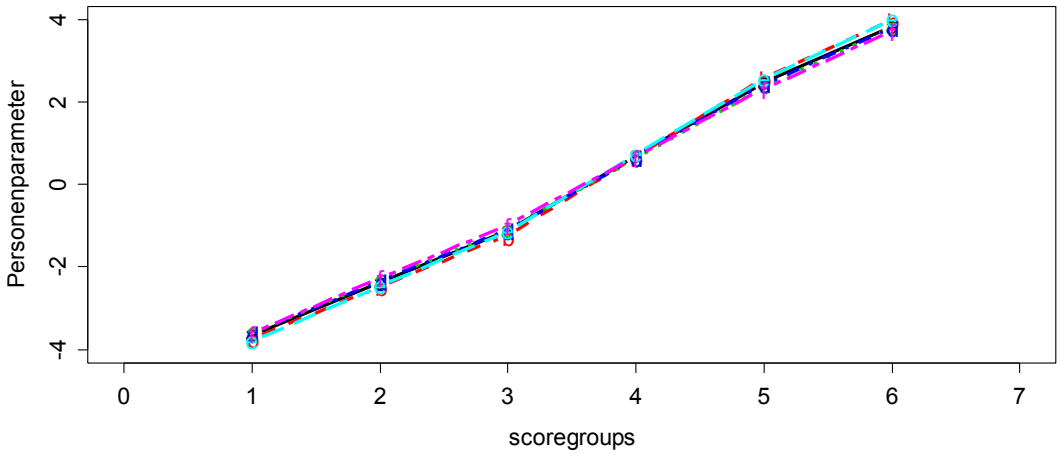
Linien = Datensätze 1 bis 6
Missinganteil 5% – 35%



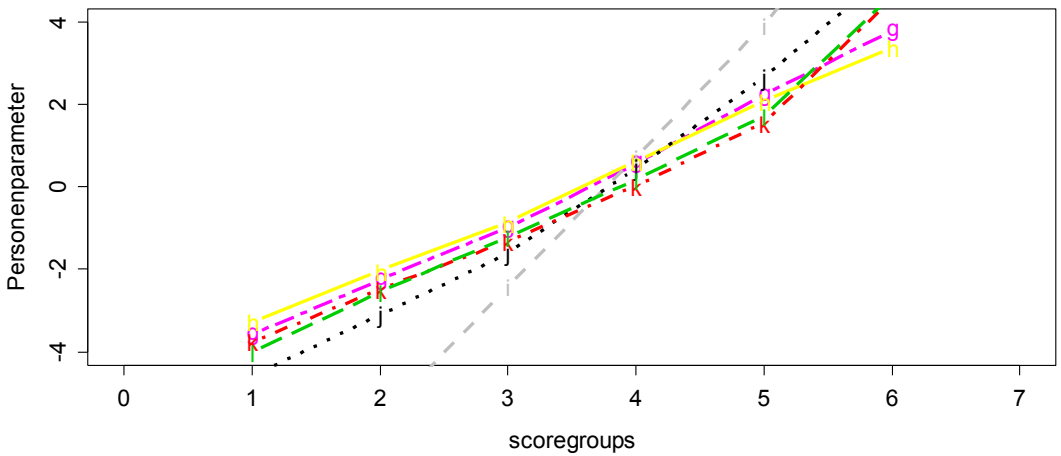
Linien = Datensätze 7 bis 12
Missinganteil 40% – 60%

Parameterbestimmung Simulation WinMira (cML)

2. Personen Parameter



Linien = Datensätze 1 bis 6
Missinganteil 5% – 35%



Linien = Datensätze 7 bis 12
Missinganteil 40% – 60%

Fazit / Ausblick

- Parameterschätzung Empirie
  - Beide Methoden (cML vs. pairwise) resultieren in vergleichbaren Item- und Personenparametern
- Parameterschätzung Simulation:
  - Item Parameter mit „pairwise comparison“ auch bei höheren missing Anteilen stabil bestimmbar
  - Personenparameter mit „pairwise comparison“ bei höheren missing Anteilen stabil bestimmbar
  - Item Parameter mit „WinMira (cML)“neigt bei höheren missing Anteilen zu extremen Schätzwerten für die Item parameter einzelner Items.
  - Personenparameter mit „WinMira (cML)“ bei höheren missing Anteilen vergleichsweise gute **aber verzerrte** Schätzung

- Ausblick – Things to do ...
  - Generalisierung / Programmierung des pairwise Algorithmus für Polytome Item-Formate ✔
  - Programmierung / Implementierung eines Modelltests ★
  - Programmierung / Implementierung Berechnung von Standardfehlern für die Itemparameter ★

Literatur

Choppin, B. (1968). Item Bank using Sample-free Calibration. *Nature*, *219*(5156), 870-872.

Choppin, B. (1985). A fully conditional estimation procedure for Rasch model parameters. *Evaluation in Education*, *9*(1), 29-42.

Davier, M. v. (2001). Winmira 2001. Groningen, The Netherlands: ASC-Assessment Systems Corporation USA and Science Plus Group.

Garner, M., & Engelhard, G., Jr. (2009). Using paired comparison matrices to estimate parameters of the partial credit Rasch measurement model for rater-mediated assessments. *Journal of Applied Measurement*, *10*(1), 30-41.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174.

Rasch, G. (1980). *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago & London: The University of Chicago Press.

R Development Core Team. (2011). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Abgerufen von http://www.R-project.org/

Zwinderman, A. H. (1995). Pairwise Parameter Estimation in Rasch Models. *Applied Psychological Measurement*, *19*(4), 369-375.