A new statistical method for detecting Differential Item Functioning in the Rasch-Model

Julia Kopf
Achim Zeileis
Carolin Strobl
LMU

Gist of this presentation

Main idea:

- Method to detect parameter instability in the Rasch-model
- Usage of model-based recursive partitioning algorithm
- Application of the method to detect DIF

Surprising result:
Higher general knowledge in Rhineland-Palatinate comparing to other German Federal states

Surprising result:

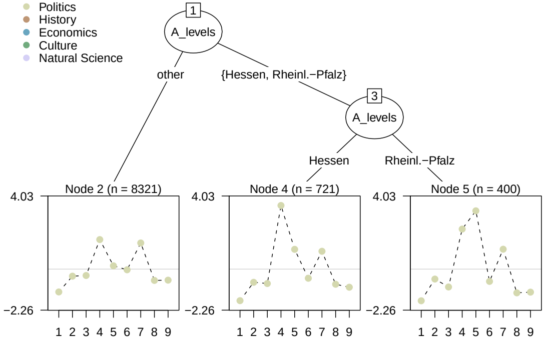
Higher general knowledge in Rhineland-Palatinate comparing to other German Federal states

- Possible reason:
  Differential Item Functioning in the Rasch-Model

Surprising result:

Higher general knowledge in Rhineland-Palatinate comparing to other German Federal states

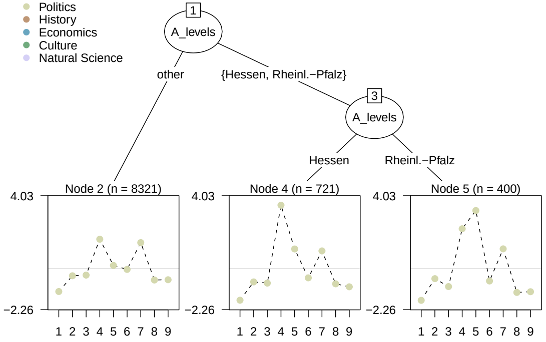
- Possible reason:
  Differential Item Functioning in the Rasch-Model



Surprising result:

Higher general knowledge in Rhineland-Palatinate comparing to other German Federal states

- Possible reason:
  Differential Item Functioning in the Rasch-Model

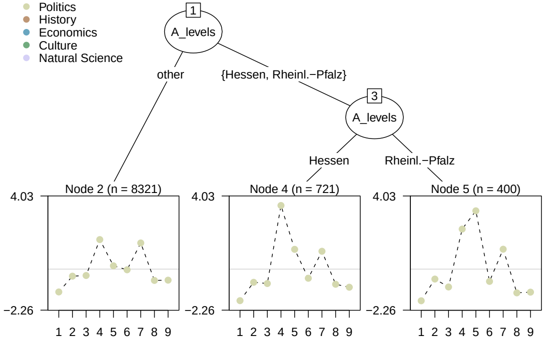


- Item 4: Find Hesse on the German map!
- Item 5: What's the capital of Rhineland-Palatinate?

Surprising result:

Higher general knowledge in Rhineland-Palatinate comparing to other German Federal states

- Possible reason:
  Differential Item Functioning in the Rasch-Model



- Item 4: Find Hesse on the German map!
- Item 5: What's the capital of Rhineland-Palatinate?
  Obtained result:

⇒ The questions in the survey do not lead to *fair comparisons*.

Objective of the Rasch-model:

- Measurement of latent variables
- Obtain at least interval scaled person parameters
- These are monotone transformation of raw scores

- Examples:
  Intelligence and attainment tests
- Extensions:
  2-pl (Birnbaum), 3-pl models

Essential data:

	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	sex*	domicile*
1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	man	west
5	0	0	0	0	1	1	0	0	0	0	0	0	0	0	woman	west
7	0	1	1	0	1	1	0	1	0	0	0	1	0	0	man	west
8	1	1	1	0	1	1	0	0	0	1	0	0	0	0	man	west
10	1	0	1	0	1	1	0	0	0	1	0	0	0	0	man	west
11	1	0	0	0	1	1	0	0	0	1	0	0	0	0	woman	west

Assumptions of the Rasch-model (Rasch, 1960):

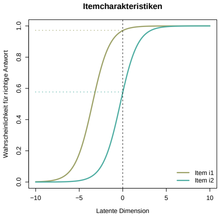
- Influence of latent variable
- Assumptions about Item Characteristic Curves (ICC)
- Unidimensionality
- Local stochastic independence
- Invariance of Item parameters

„The importance of the property of invariance of item and ability parameters cannot be overstated. This property is the cornerstone of item response theory and makes possible such important applications as equating, item banking, investigation of item bias, and adaptive testing" (Hambleton, Swaminathan and Rogers, 1991: 25).

Item Characteristic Curves and estimation process

Assumptions about the ICCs:

- Probability of solving or agreeing as a function of
  - latent variable
  - item difficulty
- monotone, logistic form



Estimation via *Conditional Maximum Likelihood* (CML):

Probability for person i $(= 1 \ldots, n)$ solving item j $(= 1 \ldots, k)$ is:

$$\mathbb{P}\left(U_{ij} = u_{ij} \,|\, \theta_i, \beta_j\right) = \frac{\exp[(\theta_i - \beta_j) \cdot u_{ij}]}{1 + \exp(\theta_i - \beta_j)},$$

- β_j denotes the item parameter of item j
- θ_i is the person parameter of individual i
- $u_{ij} \in \{0; 1\}$ symbolizes the answer of person i to item j
- $r_i = \sum_{j=1}^{k} u_{ij}$ and $s_j := \sum_{i=1}^{n} u_{ij}$

CML estimation and subject-wise score functions

New parametrisation (Fischer und Molenaar, 1995): $\xi_i = \exp(\theta_i)$ and $\varepsilon_j = \exp(-\beta_j)$

Individual Loglikelihoods:

$$\Psi(y_i, \varepsilon) = \sum_{j=1}^{k} u_{ij} \log(\varepsilon_j) - \log \gamma_{r_i}$$

with

$$\gamma_{r_i} = \sum_{\sum_{j=1}^{k} u_{ij} = r_i} \prod_{j=1}^{k} \varepsilon_j^{u_{ij}}.$$

Elementary symmetric functions (Liou, 1994):

$$\begin{aligned}
\gamma_0 &= 1 \\
\gamma_1 &= \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_k \\
\gamma_2 &= \varepsilon_1 \cdot \varepsilon_2 + \varepsilon_1 \cdot \varepsilon_3 + \ldots + \varepsilon_{k-1} \cdot \varepsilon_k \\
&\vdots \\
\gamma_k &= \varepsilon_1 \cdot \varepsilon_2 \cdot \ldots \cdot \varepsilon_k
\end{aligned}$$

Individual Scores:

$$\psi(y_i, \varepsilon^\star) = \frac{\partial \Psi(y_i, \varepsilon)}{\partial \varepsilon^\star} = \frac{u_{ij^\star}}{\varepsilon^\star} - \frac{\gamma_{r_i - 1}^{(j^\star)}}{\gamma_{r_i}}.$$

Implementation of Achim Zeileis in `psychotree`

The code conversation can be summarized in the following way:

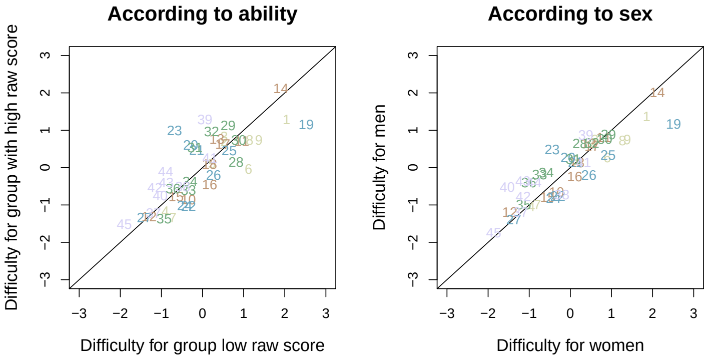
1. Hand-off formula like `item1 + item2 + ... + itemk~X1 + X2 + ... + Xl`, arguments, data
2. Model class `RaschModel` including `RaschModel.fit`
3. Data sanity checks
4. Passing to `mob()` from package `party` (Zeileis et al., 2008)

Available functions in updated package `psychotree`:

- `summary()`
- `plot()`
- `coef()`
- `worth()`

Ways of identifying violation of parameter invariance:

- Graphical model test according to item raw scores and sex



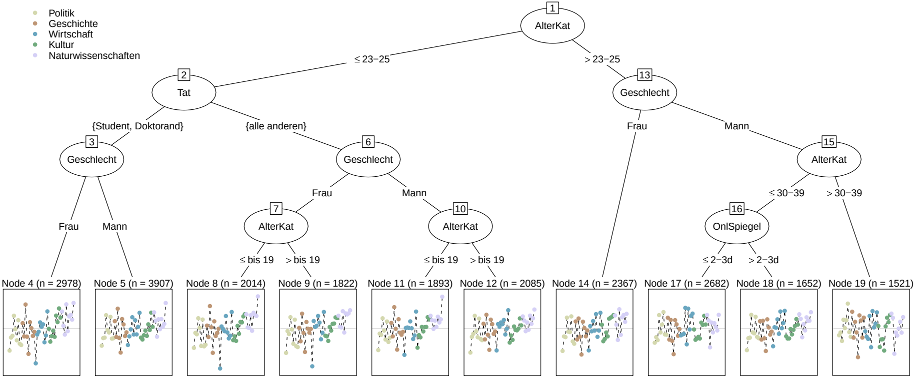
- Likelihood Ratio tests

Problem:

- Which groups may influence the item parameters?

Ways of identifying parameter variance:

New method: Rasch trees



Advantages:

- Groups are found automatically
- Statistical influence is tested
- Promising simulation results

Open questions and possible topics:

- Extended simulations, e.g. combination of covariate types
- Post-hoc tests: Which items have significant DIF?
- Extensions of Item Response Theory
- Criteria of tree stability

Fischer, G. und Molenaar, I. (1995): *Rasch Models - Foundations, Recent Developements, and Applications*. New York: Springer.

Hambleton, R., Swaminathan, H. und Rogers, H. (1991): *Fundamentals of Item Response Theory*. Newbury Park: Sage Publications.

Liou, M. (1994): *More on the Computation of Higher-Order Derivatives on the Elementary Symmetric Functions in the Rasch Model*. Applied Psychological Measurement, 18 (1), 53–62.

Rasch, G. (1960): *Probabilistic Models for some Intelligence and Attainment Tests*. Chicago, London: The University of Chicago Press.

Zeileis, A., Hothorn, T. und Hornik, K. (2008): *Model-Based Recursive Partitioning*. Journal of Computational and Graphical Statistics, 17 (2), 492–514.