On the Specification of the Score Distribution in Rasch Mixture Models

Hannah Frick, Carolin Strobl, Achim Zeileis

http://eeecon.uibk.ac.at/~frick/

Outline

- Rasch model
- Mixture models
- Rasch mixture models
- Score specifications
- Monte Carlo study
- Summary

Introduction

- Latent traits measured through probabilistic models for item response data.
- Here, Rasch model for binary items.
- Crucial assumption of measurement invariance: All items measure the latent trait in the same way for all subjects.
- Check for heterogeneity in (groups of) subjects, either based on observed covariates or unobserved latent classes.
- Mixtures of Rasch models to address heterogeneity in latent classes.

Rasch Model

Probability for person *i* to solve item *j*:

$$P(Y_{ij} = y_{ij} | \theta_i, \beta_j) \;=\; \frac{\exp\{y_{ij}(\theta_i - \beta_j)\}}{1 + \exp\{\theta_i - \beta_j\}}.$$

- y_{ij}: Response by person *i* to item *j*.
- θ_i: Ability of person *i*.
- β_j: Difficulty of item *j*.

By construction:

- No covariates, all information is captured by ability and difficulty.
- Both parameters θ and β are on the same scale: If $\beta_1 > \beta_2$, then item 1 is more difficult than item 2 for *all* subjects.

Central assumption of measurement invariance needs to be checked for both manifest and latent subject groups.

Rasch Model: Estimation

- Joint estimation of θ and β is inconsistent.
- Conditional ML (CML) estimation: Use factorization of the full likelihood on basis of the scores $r_i = \sum_{j=1}^{m} y_{ij}$:

$$\begin{aligned} L(\theta, \beta) &= f(y|\theta, \beta) \\ &= h(y|r, \theta, \beta) g(r|\theta, \beta) \\ &= h(y|r, \beta) g(r|\theta, \beta). \end{aligned}$$

  Estimate β from maximization of $h(y|r, \beta)$.
- Also maximizes $L(\theta, \beta)$ if $g(r|\cdot)$ is assumed to be independent of θ and β – regardless of the particular specification, potentially depending on auxiliary parameters δ: $g(r|\delta)$.

Mixture Model

- Assumption: Data stems from different classes but class membership is unknown.
- Modeling tool: Mixture models.
- Mixture model = $\sum$ weight $\times$ component.
- Components represent the latent classes. They are densities or (regression) models.
- Weights are a priori probabilities for the components/classes, treated either as parameters or modeled through concomitant variables.

Rasch Mixture Model: Framework

Full mixture:

- Weights: Either (non-parametric) prior probabilities π_k or weights $\pi(k|x, \alpha)$ based on concomitant variables x, e.g., a multinomial logit model.
- Components: Conditional likelihood for item parameters and specification of score probabilities

$$f(y|\alpha, \beta, \delta) = \prod_{i=1}^{n} \sum_{k=1}^{K} \pi(k|x_i, \alpha) \; h(y_i|r_i, \beta_k) \; g(r_i|\delta_k).$$

- Estimation of all parameters via ML through the EM algorithm.

Rasch Mixture Model: Estimation

- Rasch model (1 component):
  CML estimation of β independent of score specification.
- Rasch mixture model (2+ components):
  Mixture weights (also) depend on score specification. Hence, CML estimation of β also depends on the score specification.
- Unless: Score specification equal across all components.

Score Models

- Original proposition by Rost (1990): Saturated model. Discrete distribution with parameters (probabilities) $g(r) = \Psi_r$.
- Number of parameters necessary is potentially very high: (number of items $-$ 1) $\times$ (number of components).
- More parsimonious: Assume parametric model on score probabilities, e.g., using mean and variance parameters.
- Restricted score distribution: Distribution of full/unweighted sample used for each component. Estimation of β and clusters invariant to specific form.

Score Models: Intuitions

Saturated score model:

- Can capture all score distributions, i.e., never misspecified.
- Needs many (nuisance) parameters, i.e., challenging in model estimation/selection.

Mean-variance score model:

- Parsimonious, i.e., convenient for model estimation/selection.
- Potentially misspecified, e.g., for multi-modal distributions.

Restricted score model:

- Parsimonious, i.e., convenient for model estimation/selection.
- Invariant against latent structure in score distribution.
- Partially misspecified, if latent structure in scores and items coincides.

Monte Carlo Study

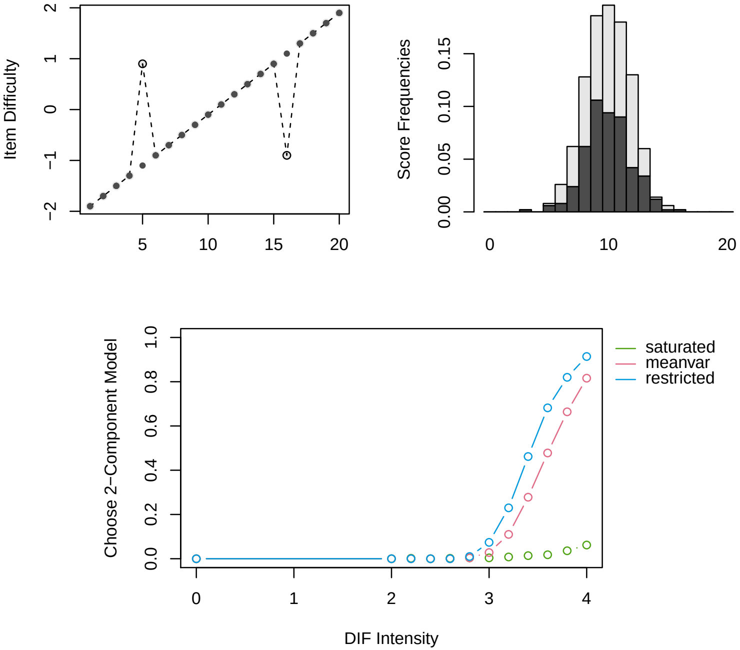
Data generating process

- 500 observations, 20 items.
- Ability: Mixture of two point masses. Difference between the two points varies from 0 to 4.
  $\rightarrow$ Resulting raw score distribution is multi-modal – or not.
- Difficulties: 2 sets with differences in 2 items, varying from 0 to 4.
  $\rightarrow$ Differential item functioning – or not.
- Grouping structure in abilities and difficulties coincides – or not.

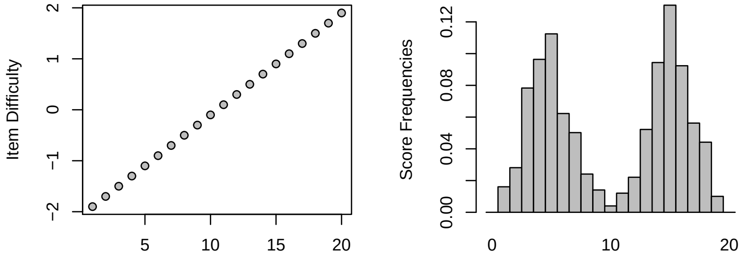
Simulation

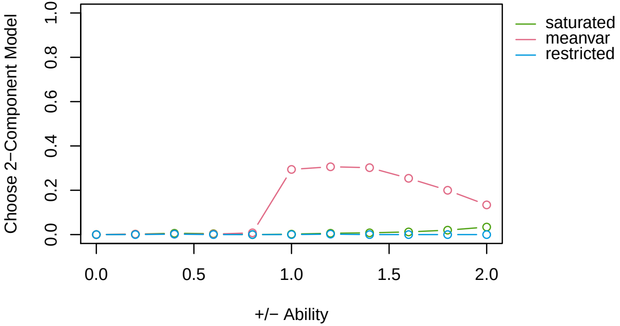
- 500 replications.
- Model fitting: various score models, several numbers of components.
- Model selection via BIC.

Differences only in Difficulties

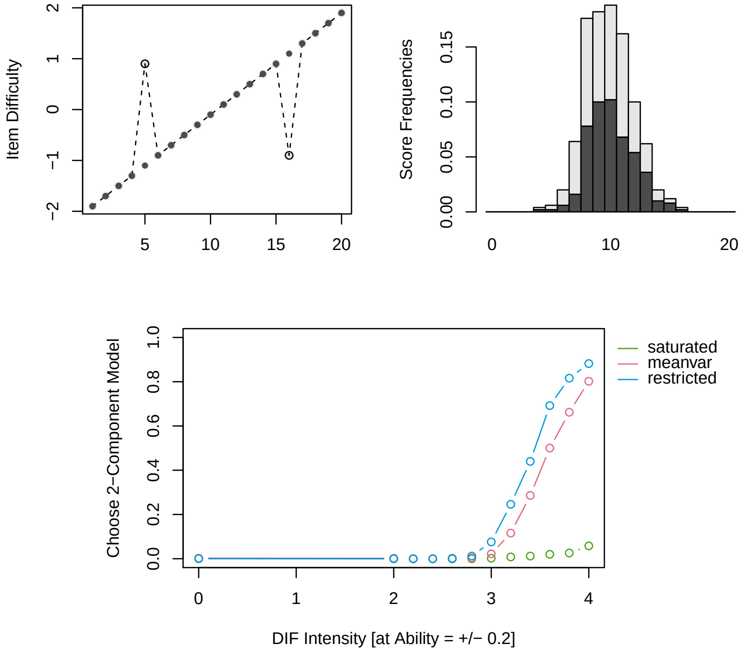


Differences only in Abilities

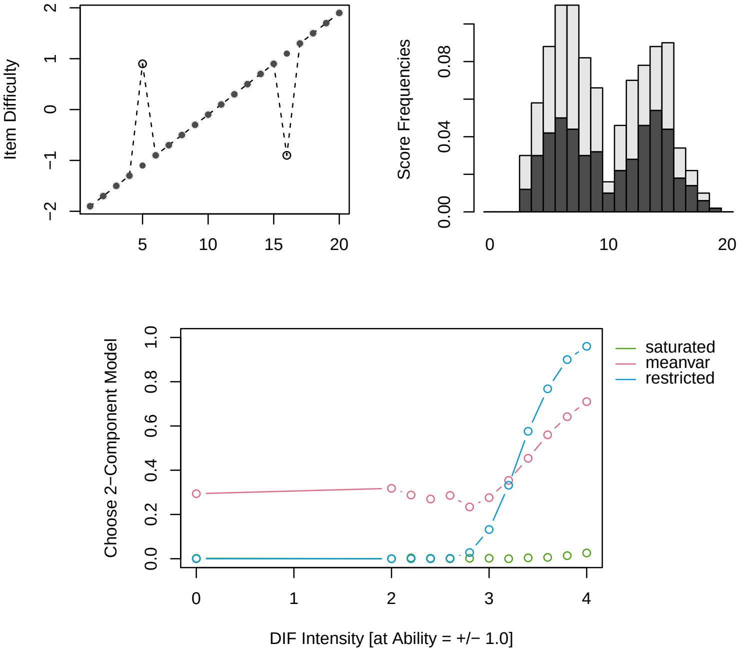




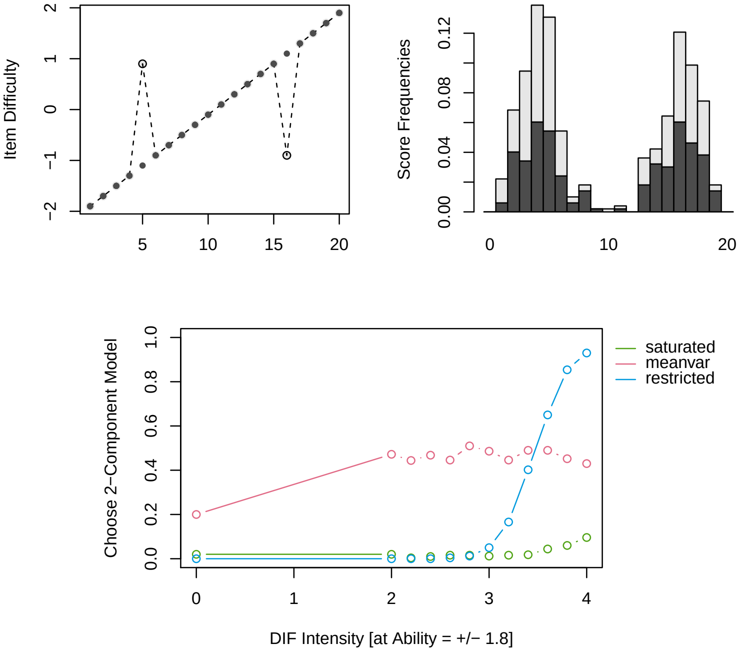
Differences in Abilities and Difficulties



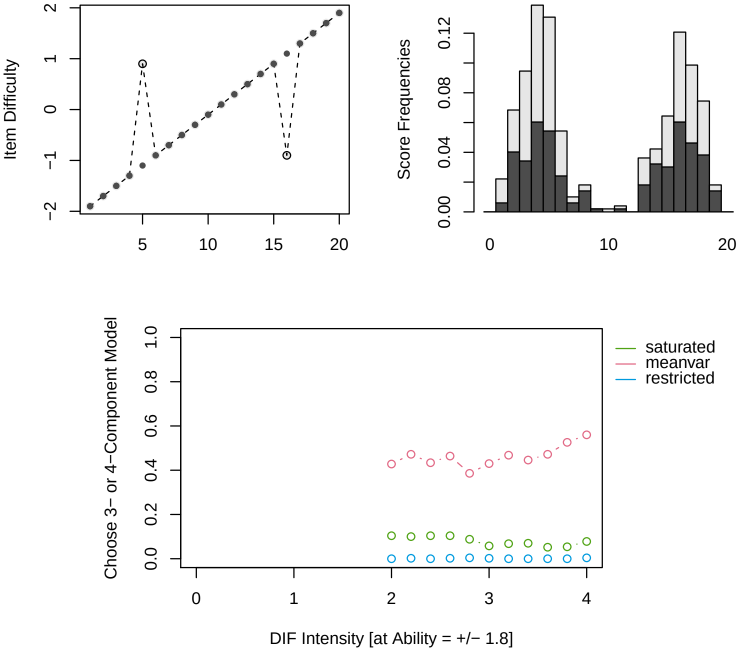
Differences in Abilities and Difficulties

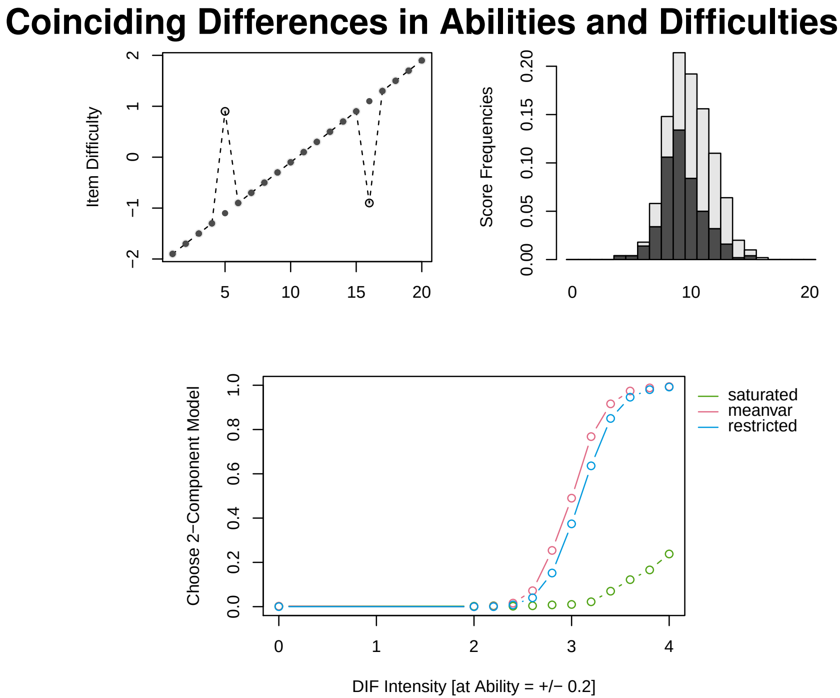


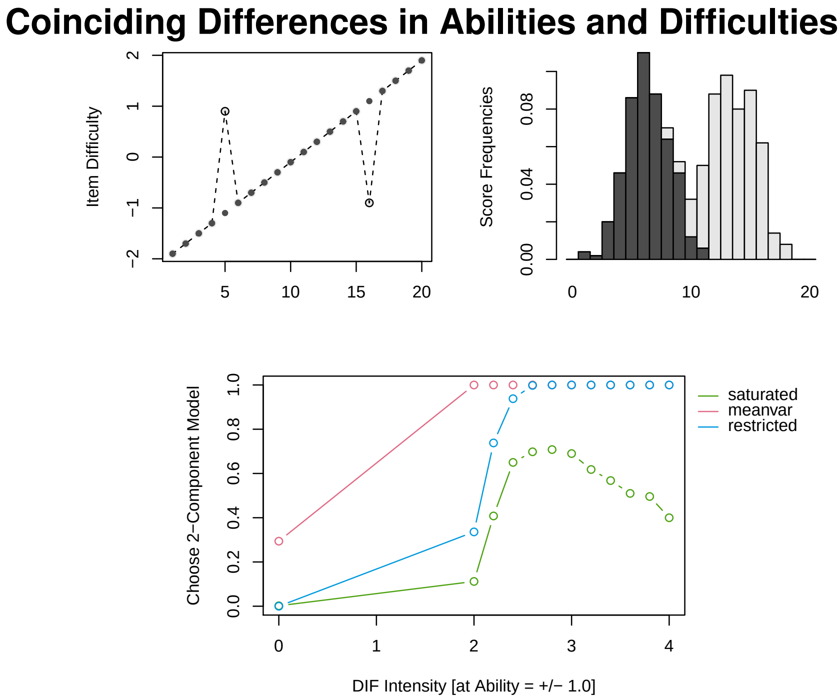
Differences in Abilities and Difficulties



Differences in Abilities and Difficulties

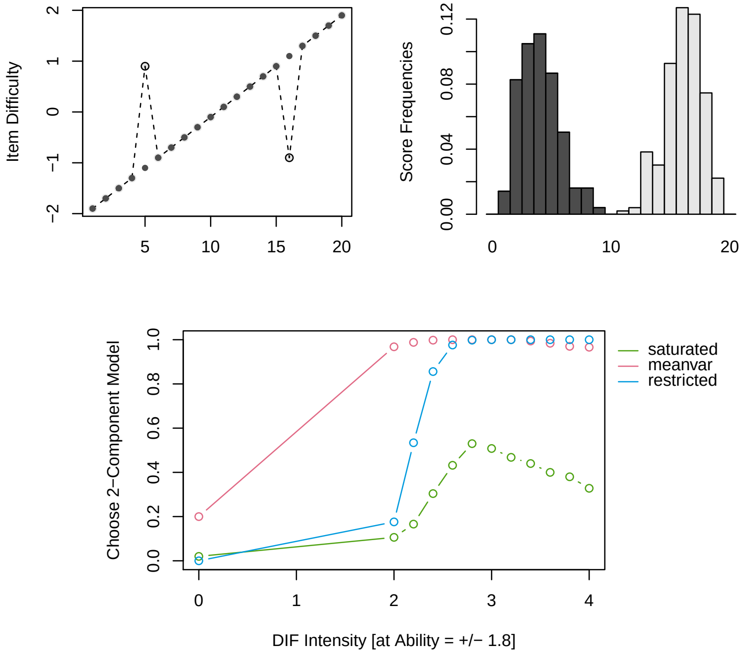



Coinciding Differences in Abilities and Difficulties
Item Difficulty
Score Frequencies
Choose 2−Component Model
saturated
meanvar
restricted
DIF Intensity [at Ability = +/− 0.2]

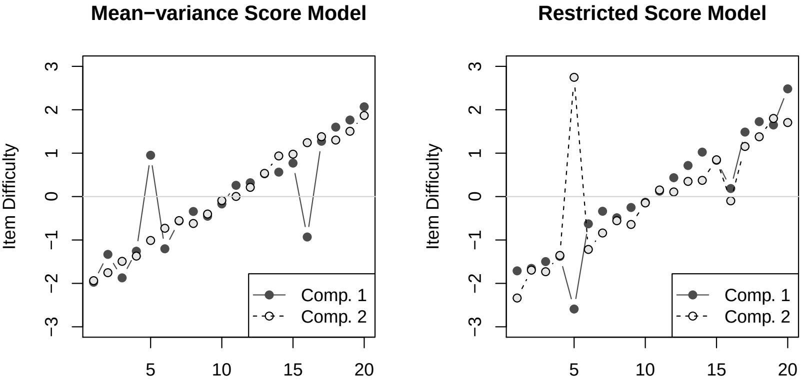


Coinciding Differences in Abilities and Difficulties
Item Difficulty
Score Frequencies
Choose 2−Component Model
saturated
meanvar
restricted
DIF Intensity [at Ability = +/− 1.0]

Coinciding Differences in Abilities and Difficulties



Coinciding Differences in Abilities and Difficulties



score model	BIC	cluster 1	cluster 2
mean-variance	11216.46	256	244
restricted	11408.67	251	249

The LR test yields a test statistic of 204.64 ($p < 0.001$).

Summary

- Rasch mixture models are a flexible means to check for measurement invariance.
- General framework incorporates various score models: saturated or mean-variance specification, possibly restricted to be equal across components.
- Restricted score distributions seem more suitable to detect the number of latent classes with regard to the item difficulties but may fail to estimate the item parameters correctly.
- Suggestion: Employ restricted score distributions to estimate the number of components. Given the number of components, compare model fit for restricted vs. unrestricted score model to choose final model.
- Implementation of all flavors soon available in R package **psychomix** at `http://CRAN.R-project.org/package=psychomix`

References

Frick H, Strobl C, Leisch F, Zeileis A (2012). “Flexible Rasch Mixture Models with Package psychomix.” *Journal of Statistical Software*, 48(7), 1–25. `http://www.jstatsoft.org/v48/i07/`

Fischer GH, Molenaar IW (eds.) (1995). *Rasch Models: Foundations, Recent Developments, and Applications*. Springer-Verlag, New York.

Grün B, Leisch F (2008). "FlexMix Version 2: Finite Mixtures with Concomitant Variables and Varying and Constant Parameters." *Journal of Statistical Software*, 28(4), 1–35. `http://www.jstatsoft.org/v28/i04/`

Rost J (1990). “Rasch Models in Latent Classes: An Integration of Two Approaches to Item Analysis.” *Applied Psychological Measurement*, 14(3), 271–282.