Universität Innsbruck

Latent Classes of Latent Traits: Mixture Models and Item Response Theory

Hannah Frick, Carolin Strobl, Friedrich Leisch, Achim Zeileis

http://www.uibk.ac.at/statistics/

Outline

- Latent traits
- Rasch model
- Latent classes and mixture models
- Rasch mixture models
- Application: Verbal aggression
- Software

Latent traits

- Aim: Measure latent traits.
- Examples:
  - Intelligence, abilities (e.g., knowledge, teamwork).
  - Attitudes (e.g., towards strangers, the EU).
  - Responsiveness to advertising.
  - Altruism, ...
- Measurement tool: Sets of items, e.g., problem solving for measuring ability, agreement with statements for measuring attitudes.
- Here: Binary items. Solve a problem y/n, agree with a statement y/n.
- State of the art model for binary items in item response theory: Rasch model.

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.
- Difference between abilty and difficulty drives probability.
- 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.

Assumptions of the Rasch model

- Central assumption: Measurement invariance.
- Violated if an item is more difficult for some groups of subjects than others.
- Called differential item functioning (DIF).
- No fair comparisons between subjects are possible based on items with DIF.
- Check for DIF before employing a Rasch model to measure a latent trait.
- Groups for which DIF occurs may be
  - covered by covariates, e.g. gender, or
  - latent (i.e. in scale) and not accessible via covariates.

Latent classes and mixture models

- 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.

Rasch mixture models: Components

- Joint estimation of θ and β is inconsitent.
- Conditional ML 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.
- However, for a mixture of Rasch models, some distribution $g(r|\cdot)$ for the score probabilities needs to be assumed, even if independent of θ and β.

Rasch mixture models: Score probabilities

- Original proposition by Rost (1990): 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 parameteric model on score probabilities, e.g., using mean and variance parameters.
- General approach: Conditional logit model encompassing the original saturated parameterization and a mean/variance parametrization (with only two parameters per component) as special cases

$$g(r|\delta) \; = \; \frac{\exp\{z_r^\top \delta\}}{\sum_{j=1}^{m-1} \exp\{z_j^\top \delta\}}.$$

Rasch mixture models

Component weights:

- Simple (non-parametric) prior probabilities π_k for each class.
- Weights $\pi(x, \alpha_k)$ based on concomitant variables x, e.g., a multinomial logit model.

Full mixture:

- Weights: With or without concomitant variables.
- Components: Conditional likelihood for item parameters and specification of score probabilities

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

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

Verbal aggression: Data

- Behavioral study of psychology students: 243 women and 73 men.
- Description of frustrating situations:
  - S1: A bus fails to stop for me.
  - S2: I miss a train because a clerk gave me faulty information.
- Behavioral mode: Want or do.
- Verbally aggressive response: Curse, scold, or shout.
- 12 resulting items: S1WantCurse, S1DoCurse, S1WantScold, . . . , S2WantShout, S2DoShout
- Covariates: Gender and an anger score.

Verbal aggression: Model selection

- Score probabilities: Mean/variance specification.
- Concomitants: With or without both covariates.
- Components: 1 to 4.
- Model choice based on BIC: 3 components without concomitants.

Number of components	1	2	3	4
Without concomitants	3874.6	3857.6	*3854.4*	3887.4
With concomitants	–	3859.1	3854.8	3880.5

Table: BIC for various Rasch mixture model specifications.

Verbal aggression: Rootogram

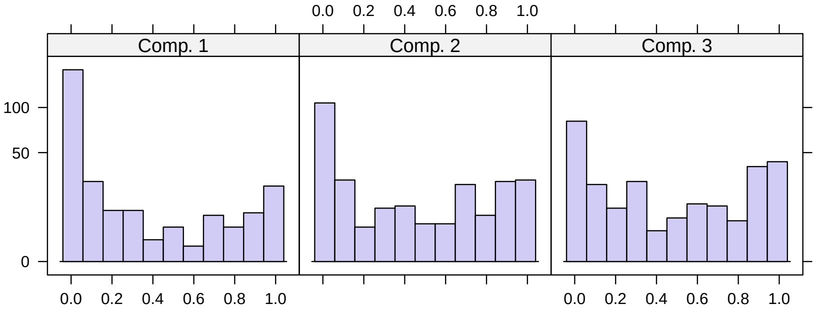

Figure: Rootogram of posterior probabilties in the 3-component Rasch mixture model.

Verbal aggression: Item profiles

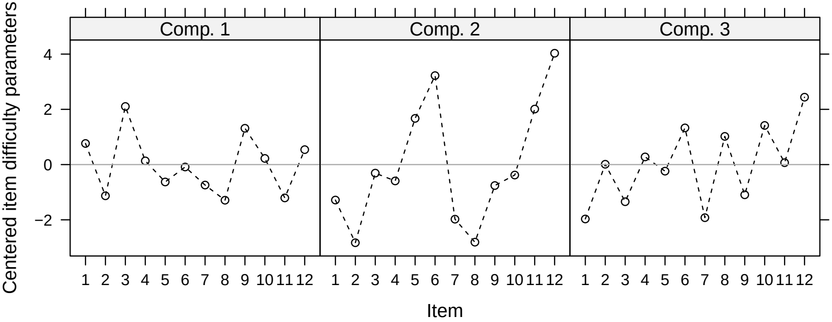

Figure: Item difficulty profiles for the 3-component Rasch mixture model. Items 1–6: Situation S1 (bus). Items 7–12: Situation S2 (train). Order: want/do curse, want/do scold, want/do shout.

Verbal aggression: Summary

- Number of components: 3 different sets of item parameters necessary.
- Not closely linked to covariates (gender, anger score) because of poorer BIC compared to model without covariates.
- Relationship between items differs between the latent classes.
- For shouting: Want is less extreme than do. For cursing and scolding, this depends on the latent class.
- One class does not differentiate much between the items, for the two other classes, cursing/scolding/shouting is increasingly extreme.

Software

- Available in R in package **psychomix** at `http://CRAN.R-project.org/package=psychomix`
- Based on package **flexmix** (Grün and Leisch, 2008) for flexible estimation of mixture models.
- Based on package **psychotools** for estimation of Rasch models.
- Frick et al. (2011), provides implementation details and hands-on practical guidance. See also `vignette("raschmix", package = "psychomix")`.

References

Frick H, Strobl C, Leisch F, Zeileis A (2011). "Flexible Rasch Mixture Models with Package psychomix." *Working Paper 2011-21*, Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, Universität Innsbruck.
`http://EconPapers.RePEc.org/RePEc:inn:wpaper:2011-21`

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 Psycological Measurement*, 14(3), 271–282.