eRm – extended Rasch modeling

An R Package for the Application of IRT Models

Hatzinger, R., & Mair, P.[1]

[1] Department of Statistics and Mathematics,
Vienna University of Economics

Item Response Theory (IRT) – Rasch Model

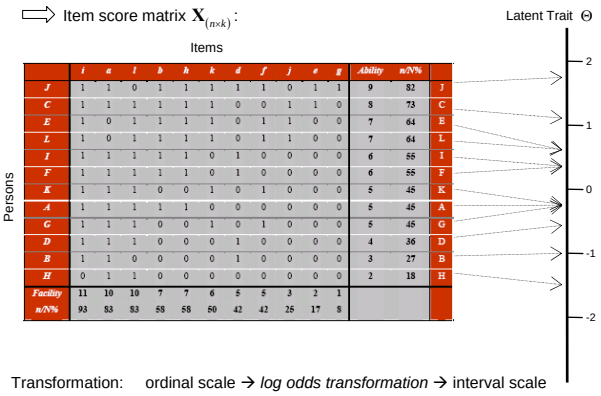
The probability of a person's expected response to an item in a test is the joint function of that person's ability (location on a latent trait) and some parameters characterizing the item.

A simple and widely used model is the **Rasch model** (Rasch, 1960) with only one parameter per item:

$$p(X_{vi} = x_{vi} \mid \theta_v, \beta_i) = \frac{\exp\left[x_{vi}(\theta_v - \beta_i)\right]}{1 + \exp(\theta_v - \beta_i)}$$

- Item parameters („difficulty"): $\beta_1, \ldots, \beta_i, \ldots, \beta_k$
- Person parameters („ability"): $\theta_1, \ldots, \theta_v, \ldots \theta_n$
- Dichotomous item score: $x_{vi} \in \{0,1\}$

⇨ Item score matrix $\mathbf{X}_{(n \times k)}$:



	I	A	L	B	H	K	D	F	J	E	Ability	n/N%	
J	1	1	0	1	1	1	1	1	0	1	9	82	J
C	1	1	1	1	1	1	0	0	1	1	8	73	C
E	1	0	1	1	1	1	0	1	1	0	7	64	E
L	1	0	1	1	1	1	0	1	1	0	7	64	L
I	1	1	1	1	1	0	1	0	0	0	6	55	I
F	1	1	1	1	1	0	1	0	0	0	6	55	F
K	1	1	1	0	0	1	0	1	0	0	5	45	K
A	1	1	1	1	1	0	0	0	0	0	5	45	A
G	1	1	1	0	0	1	0	1	0	0	5	45	G
D	1	1	1	0	0	0	1	0	0	0	4	36	D
B	1	1	0	0	0	0	1	0	0	0	3	27	B
H	0	1	1	0	0	0	0	0	0	0	2	18	H
Facility	11	10	10	7	7	6	5	5	3	2	1		
n/N%	93	83	83	58	58	50	42	42	25	17	8		

Transformation: ordinal scale → *log odds transformation* → interval scale

Estimation: $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\theta}}$ by using conditional maximum likelihood (CML)

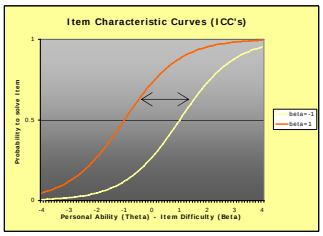
Main Assumptions of the Rasch Model

• **Unidimensionality**: The individuals performance in a test depends on a *single* underlying trait. In this case, person ability and item difficulty can be estimated and these values lie on the same interval scale. This assumption implies that the ICC's (Item characteristic curves) are invariant across subpopulations.

• **Raw Score Sufficiency:** The person and item raw score are sufficient statistics for the response pattern.

• **Parallel ICC's:** The probability of a response depends only on one latent ability parameter. The ICC's differ only in location.

• **Local Independence:** An item response does not depend on responses of the remaining items.



Model fit: Are the data consistent with the Rasch assumptions above?

$$LR = 2\left(\sum_{g=1}^{G} \log L_C\left(\hat{\boldsymbol{\beta}}_g; \mathbf{X}_g\right) - \log L_C\left(\boldsymbol{\beta}; \mathbf{X}\right)\right)$$

Linear Logistic Test Model (LLTM)

• Linear Decomposition of the item parameters.

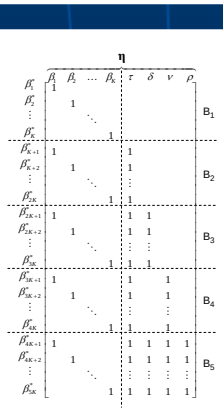
$$\boldsymbol{\beta} = \mathbf{W}\boldsymbol{\eta}$$

• Concept of "virtual items" for measuring change and comparing groups.

Example: Test with measurement points T_1, T_2

- Trend
- Gender effect
- Treatment effect
- Treatment x gender interaction

⇨ τ, δ, ν, ρ



B_1 … all subjects; T_1
B_2 … control group, males; T_2
B_3 … control group, females; T_2
B_4 … treatment group, males; T_2
B_5 … treatment group, females; T_2

Linear Rating Scale Model (RSM and LRSM)

• Extension of the Rasch model for rating scales.

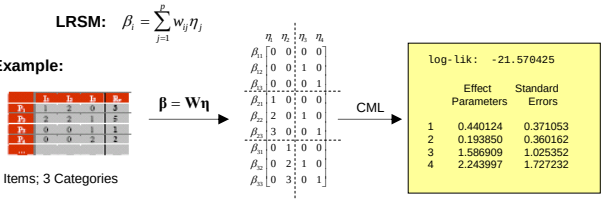
• Each item has the same number of ordered categories.

RSM: $$P(X_{vi} = k \mid \theta_v, \beta_i, \omega_0, \ldots, \omega_m) = \frac{\exp\left(k(\theta_v + \beta_i) + \omega_k\right)}{\sum_{h=0}^{m} \exp\left(h(\theta_v + \beta_i) + \omega_h\right)}$$

ω_h … category parameter; h, k … categories $h = 0, \ldots, m$

LRSM: $\beta_i = \sum_{j=1}^{p} w_{ij}\eta_j$

Example:



3 Items; 3 Categories

```
log-lik:  -21.570425

     Effect      Standard
  Parameters      Errors
1   0.440124    0.371053
2   0.193850    0.360162
3   1.586909    1.025352
4   2.243997    1.727232
```

Linear Partial Credit Model (PCM and LPCM)

• Each item category gets a partial credit.

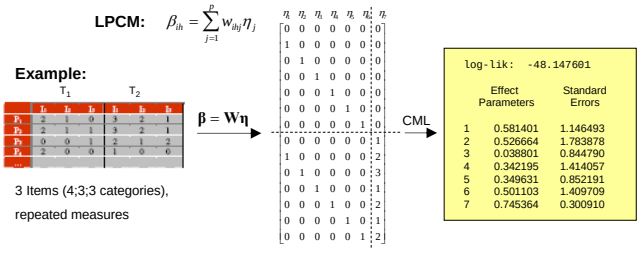
• Different number of categories per item are allowed.

PCM: $$P(X_{vik} = 1 \mid \theta_v, \beta_{ik}) = \frac{\exp\left(k\theta_v + \beta_{ik}\right)}{\sum_{h=0}^{m_i} \exp\left(h\theta_v + \beta_{ih}\right)}$$

β_{ih} … item-category parameter; h, k … categories $h = 0, \ldots, m_i$

LPCM: $\beta_{ih} = \sum_{j=1}^{p} w_{ihj}\eta_j$

Example:

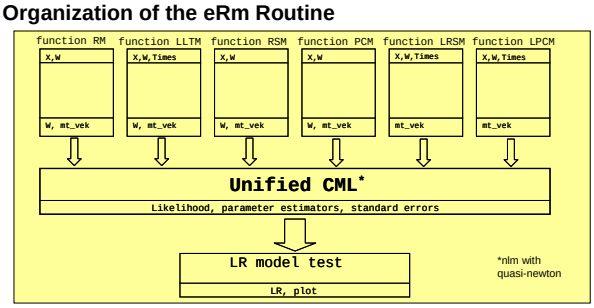


3 Items (4;3;3 categories),
repeated measures

```
log-lik:  -48.147601

     Effect      Standard
  Parameters      Errors
1   0.581401    1.146493
2   0.526664    1.783878
3   0.038801    0.844790
4   0.342195    1.414057
5   0.349631    0.852191
6   0.501103    1.409709
7   0.745364    0.300910
```

Organization of the eRm Routine



References

Fischer, G.H., & Molenaar, I. (1995). *Rasch Models: Foundations, Recent Developments, and Applications.* New York: Springer.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573.

Fischer, G.H., & Parzer, P. (1991). An extension of the rating scale model with an application to the measurement of change. *Psychometrika, 56*, 637-651.

Fischer, G.H., & Ponocny, (1994). An extension of the partial credit model with an application to the measurement of change. *Psychometrika, 59*, 177-192.

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

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: The Danish Institute of Educational Research.

Scheiblechner, H. (1972). Das Lernen und Lösen komplexer Denkaufgaben. [The learning and solving of complex reasoning items.] *Zeitschrift für Experimentelle und Angewandte Psychologie, 3*, 456-506.