Rasch Models
and the R package eRm

Rusch, Thomas; Hatzinger, Reinhold

Published: 08/04/2010

Document Version
Unknown

Rasch Models and the R package eRm

Thomas Rusch
Reinhold Hatzinger

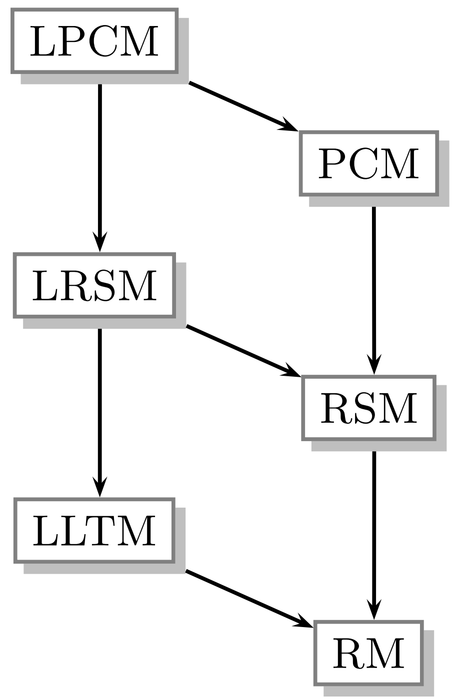
Institute for Statistics and Mathematics
WU Wien

What is eRm?

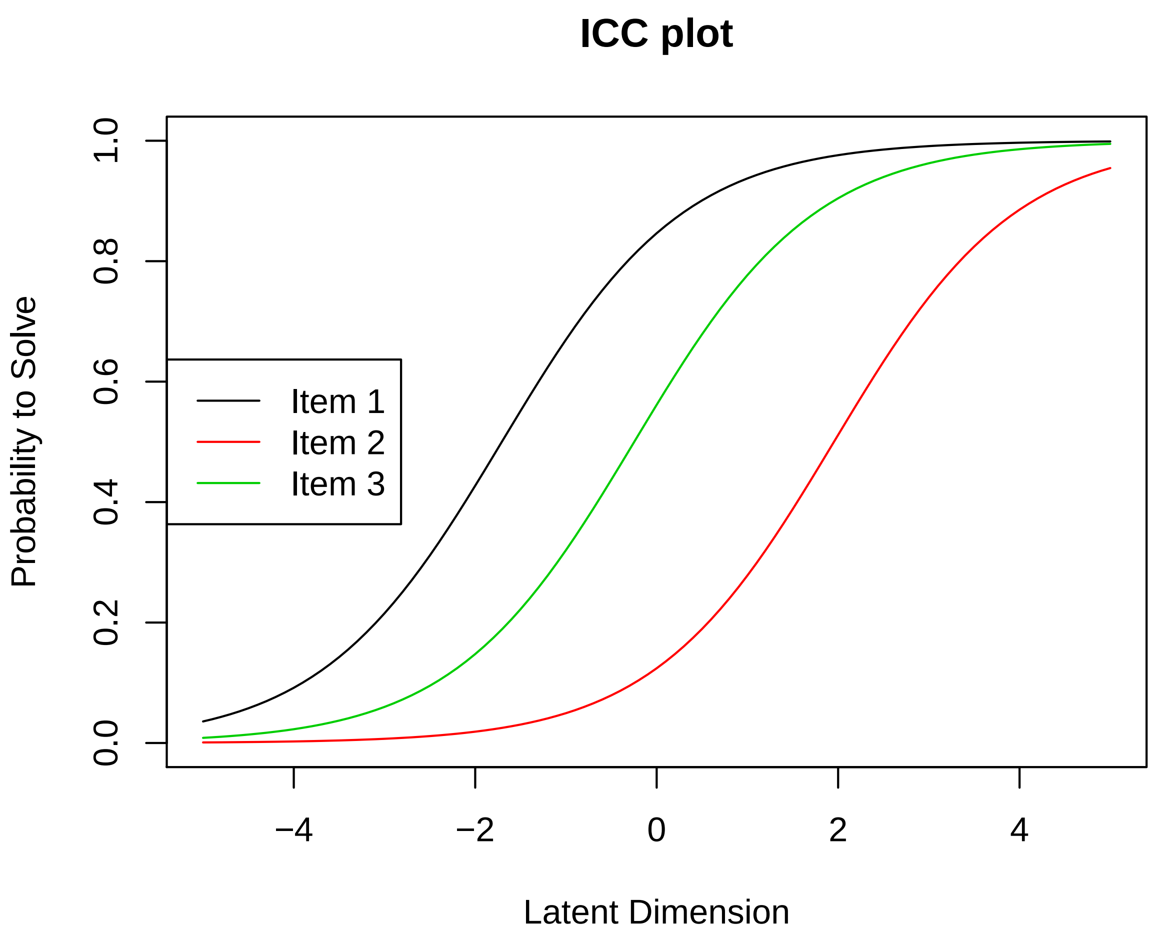
- **eRm** short for *extended Rasch modelling*
- is an **R** package
- is open source: no license fees, source code available,
  GPL: share, change, and redistribute under certain conditions
- for Rasch family models:
  utilities for fitting, testing, and displaying results
- currently implemented models:
  LPCM, PCM, LRSM, RSM, LLTM, RM, (LLRA)
- uses CML estimation

The model hierarchy in eRm

The LPCM is the most general unidimensional model in this family
All other models are submodels
They are obtained by appropriately defining the design matrix W



The Rasch Model (RM) (Rasch, 1960)



Item Parameter Estimation

Conditional Maximum Likelihood (CML) condition on r_v

$$L_c = \exp(-\sum_i \beta_i s_i) / \prod_r \sum_{x|r} \exp(-\sum_i x_i \beta_i)^{n_r}$$

– person parameters do not occur in the conditional likelihood
– items can be compared independent of persons (separation)
– leads to specific objectivity
– person free item calibration
– 'sample-independence':
actual sample not of relevance for inference on item parameters

CML estimates are unbiased and consistent as $n \to \infty$
for estimability set $\beta_1 = 0$ or $\sum \beta_i = 0$
items with score $s_i = 0$ or n and person with $r_v = 0$ or k are removed prior to estimation

Person Parameter Estimation

Weighted and unweighted ML estimation using the unconditional likelihood

$$L_u = \frac{\exp(\sum_v \theta_v r_v)\exp(-\sum_i \beta_i s_i)}{\prod_v \prod_i (1+\exp(\theta_v - \beta_i))}$$

and assuming the βs to be known (from prior estimation)

slightly biased (bias smaller than s.e.'s of estimates)
no estimates for $r_v = 0$ and $r_v = k$
can be approximated using, e.g., spline interpolation

weighted ML estimation:
likelihood function is skewed, additional source of estimation bias
Warm (1989) suggests unbiasing correction, computationally unfeasible

The R package eRm (extended Rasch modelling)

```
> library(eRm)
```

main functions concerning fit of the RM:

- `RM(data)` fits the RM and generates object of class `dRm`
- `person.parameter(drmobj)` generates object of class `ppar`
- plots from `drm` object:
  - `plotPImap()`, `plotICC()`, `plotjointICC()`
- plots from `ppar` object:
  - `plot()`
- extract information from `drm` object:
  - `coef()`, `vcov()`, `confint()`, `logLik()`, `model.matrix()`
- extract information from `ppar` object:
  - `confInt()`, `logLik()`

Fitting the RM

```
> rm.res <- RM(data)
> rm.res
Results of RM estimation:

Call:  RM(X = data)

Conditional log-likelihood: -156.3100
Number of iterations: 12
Number of parameters: 4

Item (Category) Difficulty Parameters (eta):
                 I2        I3        I4          I5
Estimate -0.4292685 1.1743542 0.1496732 -0.02667262
Std.Err   0.1945618 0.2243309 0.1918824  0.19118378
```

– default is: `RM(datamatrix, sum0 = TRUE, other options)`

– `sum0` defines constraints (for estimability):

`TRUE` ... sum zero, `FALSE` ... first item set to 0

– the output is difficulty parameters for the basis parameters in non-linearised models

```
> summary(rm.res)
Results of RM estimation:

Call:  RM(X = data)

Conditional log-likelihood: -156.3100
Number of iterations: 12
Number of parameters: 4

Item (Category) Difficulty Parameters (eta) with 0.95 CI:
   Estimate Std. Error lower CI upper CI
I2   -0.429      0.195   -0.811   -0.048
I3    1.174      0.224    0.735    1.614
I4    0.150      0.192   -0.226    0.526
I5   -0.027      0.191   -0.401    0.348

Item Easiness Parameters (beta) with 0.95 CI:
        Estimate Std. Error lower CI upper CI
beta I1    0.868      0.206    0.464    1.272
beta I2    0.429      0.195    0.048    0.811
beta I3   -1.174      0.224   -1.614   -0.735
beta I4   -0.150      0.192   -0.526    0.226
beta I5    0.027      0.191   -0.348    0.401
```

Extracting Information

the item parameter estimates

```
> coef(rm.res)
    beta I1     beta I2     beta I3     beta I4     beta I5
 0.86808629  0.42926853 -1.17435425 -0.14967319  0.02667262
```

the variance-covariance matrix of item parameter estimates

```
> vcov(rm.res)
             [,1]        [,2]         [,3]         [,4]
[1,]  0.037854298 -0.01255416 -0.008073628 -0.007959445
[2,] -0.012554163  0.05032434 -0.011716070 -0.011780076
[3,] -0.008073628 -0.01171607  0.036818873 -0.007484464
[4,] -0.007959445 -0.01178008 -0.007484464  0.036551239
```

Extracting Information (cont'd)

confindence intervals for the item parameter estimates

```
> confint(rm.res, "beta")
              2.5 %     97.5 %
beta I1  0.46444285  1.2717297
beta I2  0.04793439  0.8106027
beta I3 -1.61403470 -0.7346738
beta I4 -0.52575588  0.2264095
beta I5 -0.34804072  0.4013859
```

the conditional log likelihood

```
> logLik(rm.res)
'Conditional log Lik.' -156.3100 (df=4)
```

Person Parameter Estimation

```
> pp <- person.parameter(rm.res)
> pp
Person Parameters:

 Raw Score   Estimate Std.Error
         0 -2.6310979        NA
         1 -1.5189967 1.1498599
         2 -0.4615091 0.9565427
         3  0.4374933 0.9636305
         4  1.5217580 1.1669395
         5  2.6659917        NA
```

if NAs in the data, different person parameters are estimated for every NA-pattern group

Methods for Person Parameter Estimation Results

```
> logLik(pp)
'Unconditional (joint) log Lik.' -10.85398 (df=4)
```

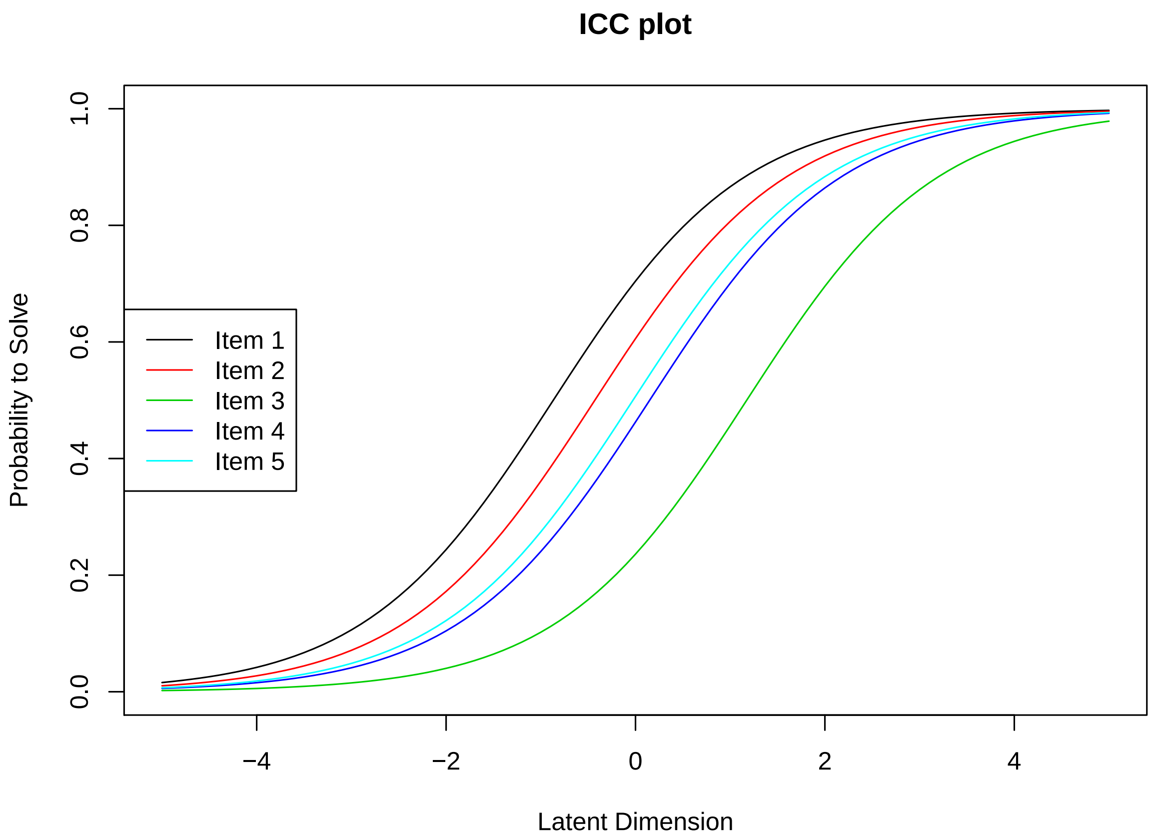
```
> confint(pp)
```

```
       2.5 %    97.5 %
P1 -3.772681 0.7346872
P2 -1.451188 2.3261743
P3 -1.451188 2.3261743
P5 -2.336298 1.4132801
P6 -1.451188 2.3261743
P7 -2.336298 1.4132801
...
```

attention: `confint(pp)` gives values for all subjects
if there are Nas in the data, confidence intervals are printed for each NA group

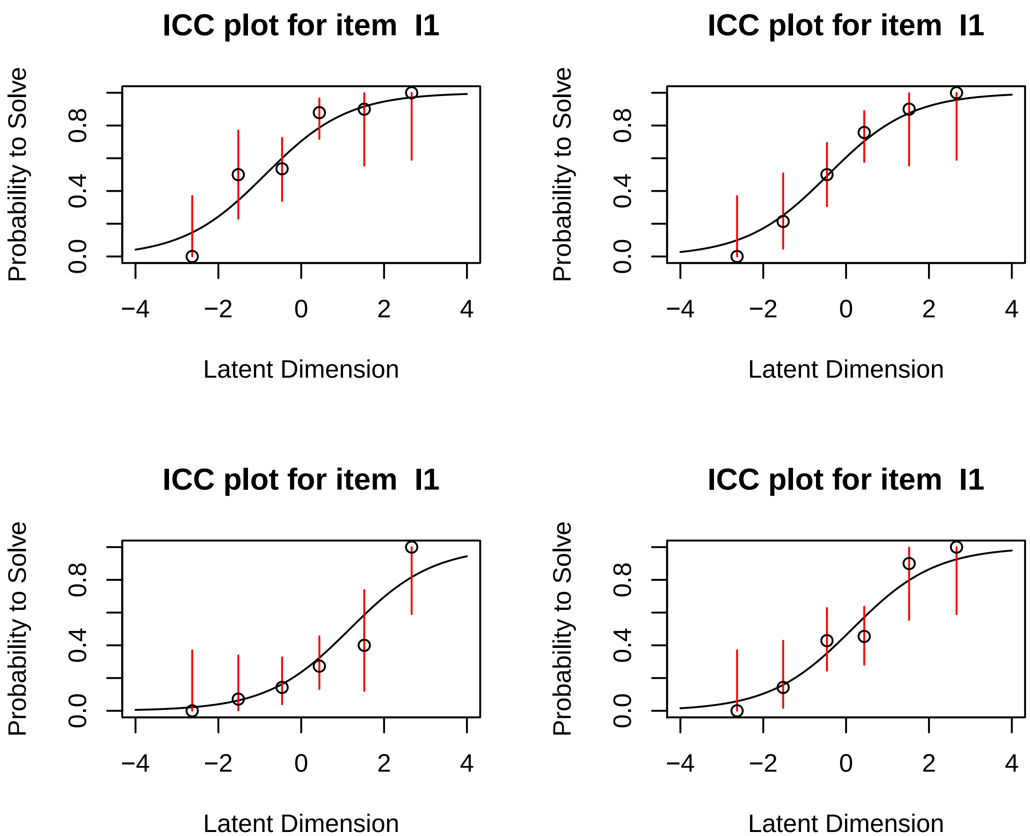
Plot ICCs

```
> plotjointICC(rm.res, xlim = c(-5, 5))
```



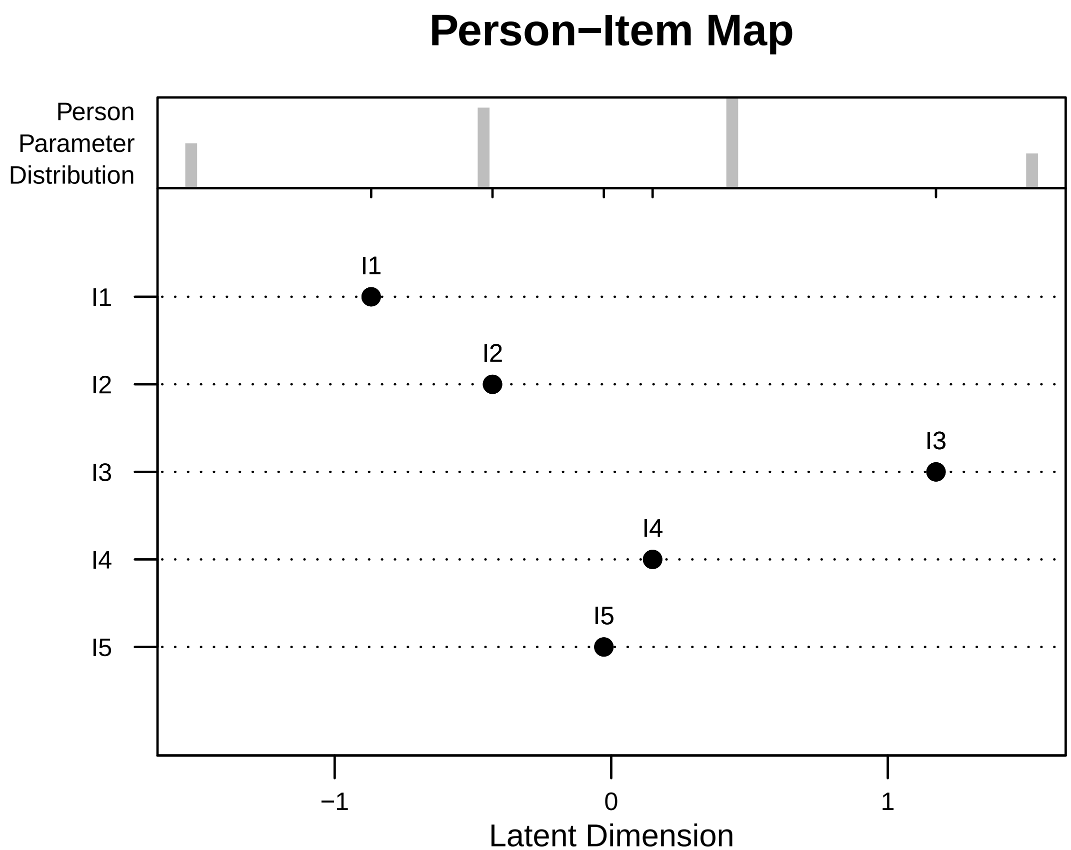
Plot ICCs

```
> plotICC(rm.res, item.subset = 1:4, ask = F, empICC = list("raw"),
+     empCI = list(lty = "solid"))
```



Plot Person-Item Map

```
> plotPImap(rm.res)
```



Plot of Person Parameter Estimates

```
> plot(pp)
```

Assessing Goodness of Fit of the RM

RM allows to evaluate the quality of measurement
crucial assumptions empirically testable
aim: find set of items that conform to the RM (‘data fit model’)

various tests/diagnostics have been proposed

some implemented in eRm:

– Andersen LR test
– Wald-type test
– nonparametric tests
– item/person fit indices
– graphical procedures
– Martin-Löf test

Andersen's Likelihood Ratio Test (Andersen, 1973)

– 'global' test (all items investigated simultaneously)
– powerful against violations of sufficiency and monotonicity
– can detect DIF (differential item functioning or *item bias*):

Wald Test

allows for testing single items idea is again: sample into subgroups (usually 2) using separate estimates $\widehat{\beta}_j^{(1)}$ and $\widehat{\beta}_j^{(2)}$

Person/Item Fit

objective is to detect noticeable patterns

Expected response: $\pi_{vi} = \exp(\theta_v - \beta_i)/(1 + \exp(\theta_v - \beta_i))$

Residuals: $e_{vi} = x_{vi} - \pi_{vi}$

Nonparametric ('exact') Tests

Idea:

- Parameter estimates depend only on marginals $\boldsymbol{r}$ and $\boldsymbol{s}$
- for any statistic of the data matrix, one can approximate the null distribution
- take random sample from the collection of equally likely data matrices, compute null distribution of statistic
- valid and powerful, even in small samples (UMP)

Statistic	Scenario
T_1,T_2,T_{11}	local dependence
T_4	DIF
T_7,T_{7a}	different discrimination
T_{10}	subgroup invariance

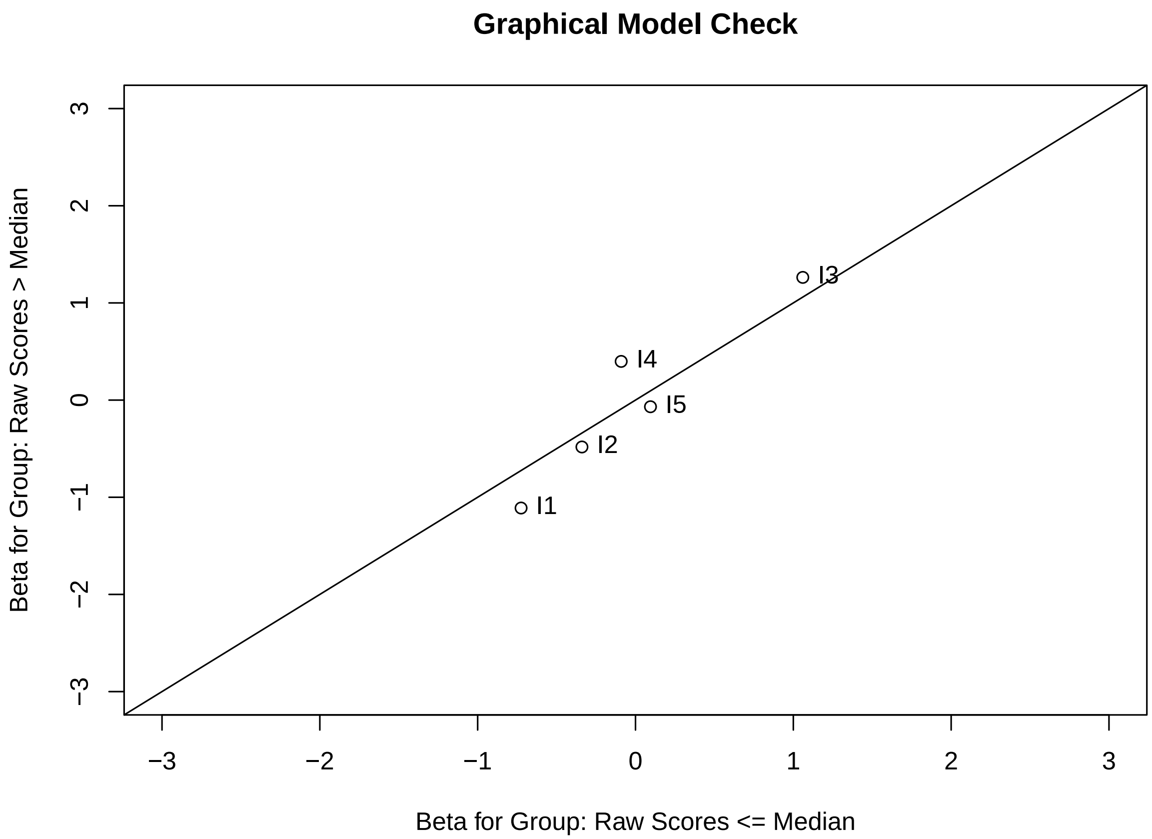
Martin-Lőf's Likelihood Ratio Test (cf. Gustafsson, 1980)

– 'global' test (all items investigated simultaneously)
– powerful against violations of unidimensionality
– can detect multidimensionality ("differential person functioning")

basic idea:
similar to Andersen's test
split *item* set in two parts, I_1, I_2
calculate the (maximum) likelihood for both parts
if Rasch Model holds the product of these likelihoods should equal the overall likelihood
Test statistic is asymptotically χ^2-distributed with $df = I_1 I_2 - 1$

Graphical Procedure

underlying idea again subgroup homogeneity, plot $\widehat{\beta}^{(1)}$ vs $\widehat{\beta}^{(2)}$



Andersen's LR Test:

```
> lrt <- LRtest(rm.res, se = TRUE)
> lrt

Andersen LR-test:
LR-value: 2.407
Chi-square df: 4
p-value:  0.661
```

Wald Test:

```
> Waldtest(rm.res)
Wald test on item level (z-values):

        z-statistic p-value
beta I1      -0.832   0.405
beta I2      -0.352   0.725
beta I3       0.428   0.668
beta I4       1.300   0.194
beta I5      -0.411   0.681
```

Item Fit:

```
> itemfit(pp)
Itemfit Statistics:
    Chisq df p-value Outfit MSQ Infit MSQ
I1 80.938 84   0.574      0.952     0.966
I2 78.491 84   0.649      0.923     0.934
I3 82.480 84   0.526      0.970     0.961
I4 85.144 84   0.445      1.002     1.024
I5 74.275 84   0.767      0.874     0.908
```

Nonparametric Tests:

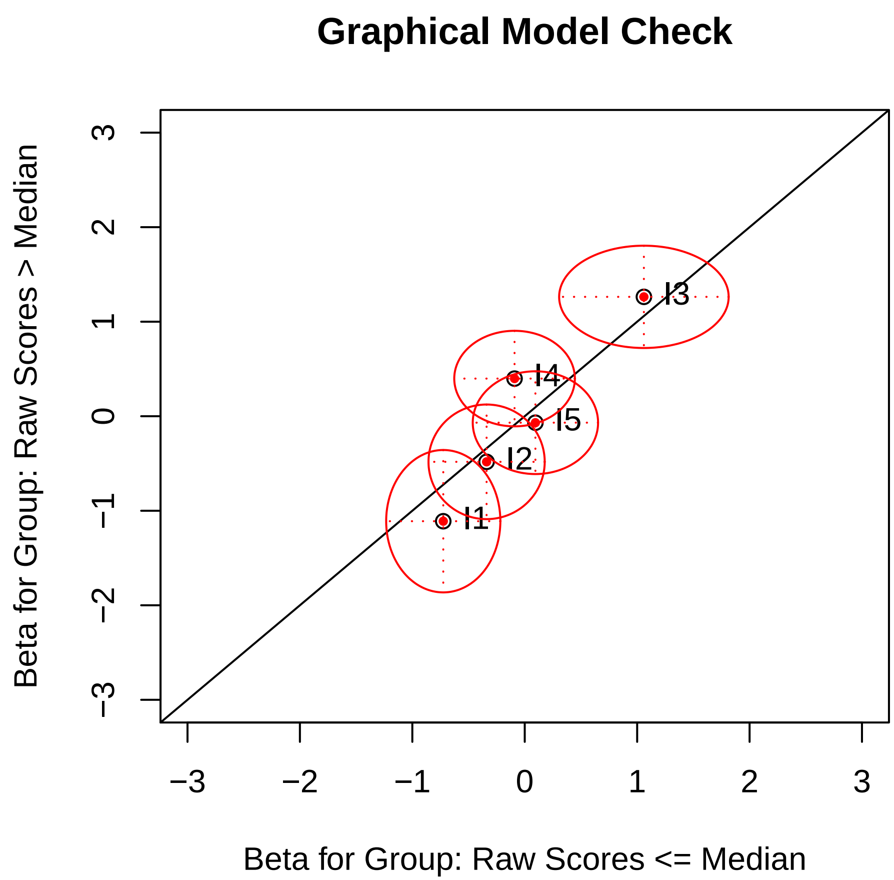
```
> t11 <- NPtest(data, method = "T11")
> t11
Nonparametric RM model test: T11 (global test - local dependence)
    (sum of deviations between observed and expected inter-item correlations)

Number of sampled matrices: 500
one-sided p-value: 0.954
```

Martin-Lőf Test:

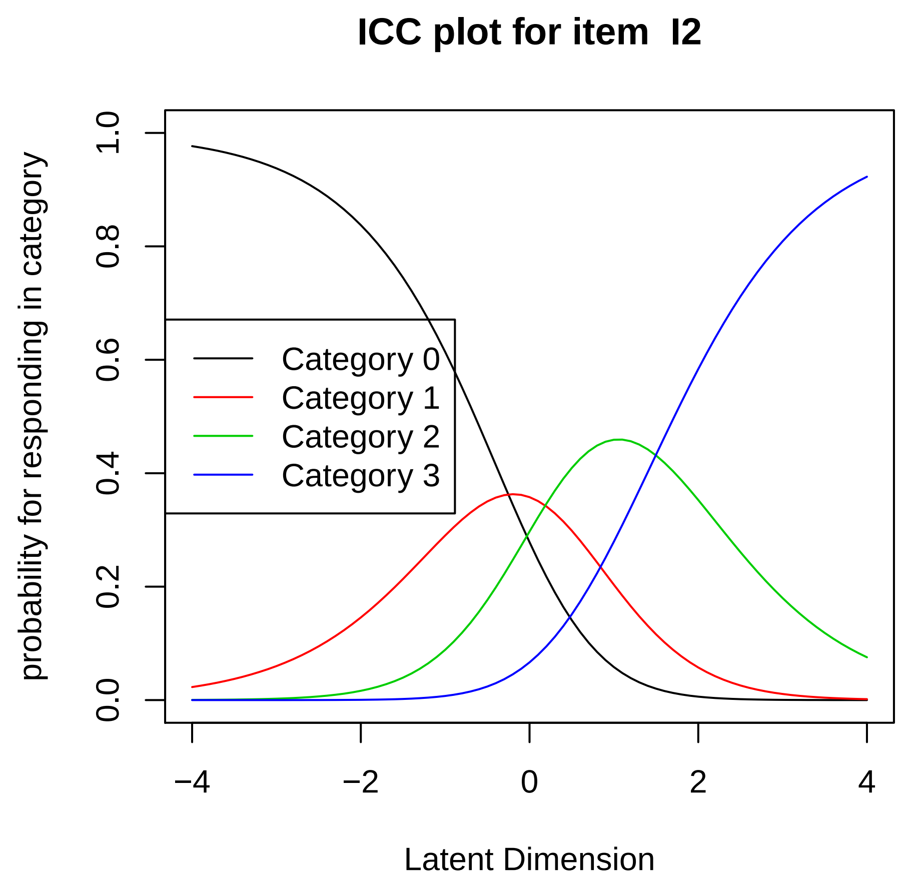
```
> MartinLoef <- MLoef(rm.res)
> MartinLoef

Martin-Loef-Test (split criterion: median)
LR-value: 4.868
Chi-square df: 5
p-value: 0.432
```

Graphical Procedure

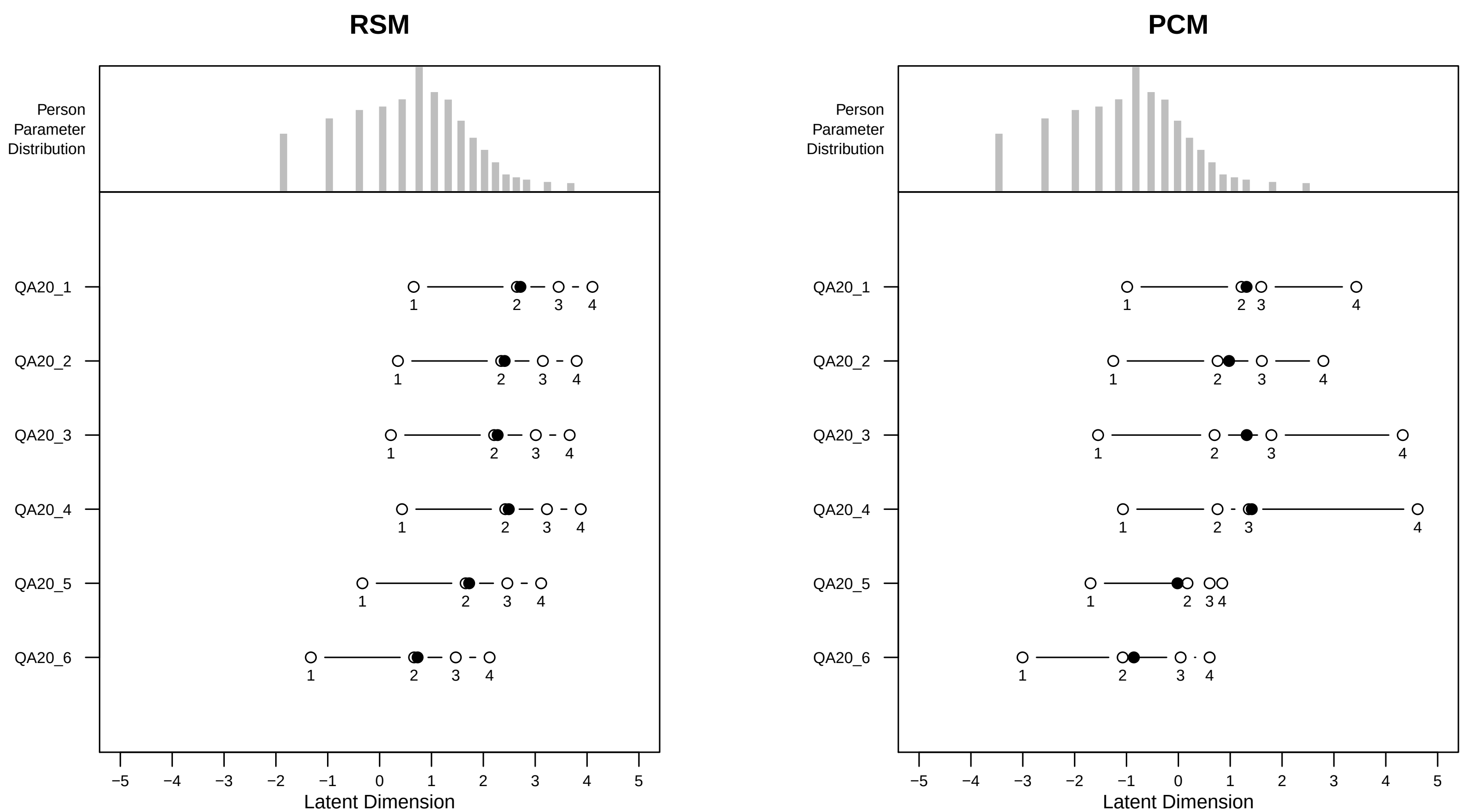
```
> plotGOF(lrt, conf = list())
```



Partial Credit Model (PCM) and Rating Scale Model (RSM)



Comparison RSM vs PCM



R commands

main functions concerning fit of polytomous models:

- `PCM(data)` fits the PCM and generates object of class `Rm`
- `RSM(data)` fits the RSM and generates object of class `Rm`
- `thresholds(rmobj)` displays the itemparameter estimates as thresholds
- all other functions are the same as previously presented (except for `plotjointICC()`)

eRm Features

Scope:

- Scale Analysis (measurement models)
- Modelling latent change (statistical models)
  uni- and multidimensional (LLRA)

Models:

- RM, RSM, PCM, LLTM, LRSM, LPCM, (LLRA)
- Treatment of missing values (MCAR)
- Different constraints for parameter estimation
- Design matrix (default / user defined)

Estimation:

- Itemparameters, 'basic'- and effect parameters,
  threshold parameters (all using CML)
- Personparameters (JML)
- Covariance matrices (confidence intervals)
- Support for stepwise item selection

eRm Summary and Features (cont'd)

Diagnostics, Model Tests, and Fit Statistics:

- Andersen LR-test, Wald Test for single items
- Global and item level nonparametric tests (for RM)
- Itemfit, Personfit (using Pearson residuals)
- Information criteria (AIC, BIC, cAIC)
- Check for existence of ML estimates – 'well-conditioned datamatrix' (for RM)
- some (nonpsychometric) logistic regression diagnostics

Plots:

- Goodness-of-Fit Plots
- ICC-Plots for single items (with optional empirical ICCs)
- Joint ICC-Plot (for RM)
- Person-Item Map

Miscellaneous :

- Simulation of data matrices according to RM violations
- . . .

Further Infos:

R Forge: http://r-forge.r-project.org/

Development platform

latest releases downloads

Discusssion and help forum

Project homepage http://erm.r-forge.r-project.org/

Publications:

Mair & Hatzinger (2007). Journal Statistical Software

Mair & and Hatzinger (2007). Psychology Science

Hatzinger & Rusch (2009). Psychology Science Quarterley