Item Fit

Rasch Technical Training 3

Carolina Fellinghauer : fellinghauerc@who.int

Item Fit

How accurately or predictably questionnaire items **fit** the Rasch model

Infit and Outfit

To find the item fit requires computation of:

1) Expected response for each observation X_{ij}

$$E_{ij} = \Sigma_{k=0}^{mi} k(P_{ikj})$$

2a) The score residual Y_{ij}

$$Y_{ij} = X_{ij} - E_{ij}$$

2b) The standardized residual Z_{ij}

$$Z_{ij} = \frac{Y_{ij}}{(W_{ij})^{1/2}}$$

The variance of X_{ij} is formalized as

$$W_{ij} = \Sigma_{k=0}^{mi} (k - E_{ij})^2 P_{ikj}$$

Infit and Outfit

3) A chi-square statistic by summing the standardized residuals.

$$\chi^2 = \Sigma_{n=1}^{N} Z_{ij}^2$$

The chi-square divided by the sample size corresponds to the Mean-Square Outfit Statistic.

$$Outfit_i = \frac{\Sigma_{n=1}^{N} Z_{ij}^2}{N}$$

The Outfit Statistic is sensitive to outlier. To diminish the effect of outlier, the standardized residuals can be adjusted by their variance. This is the Mean-Square Infit Statistic.

$$Infit_i = \frac{\Sigma_{n=1}^{N} W_{ij} Z_{ij}^2}{\Sigma_{n=1}^{N} W_{ij}}$$

Underfit and Overfit

An item is fitting the Rasch model if the Infit and Outfit statistics are close to 1.

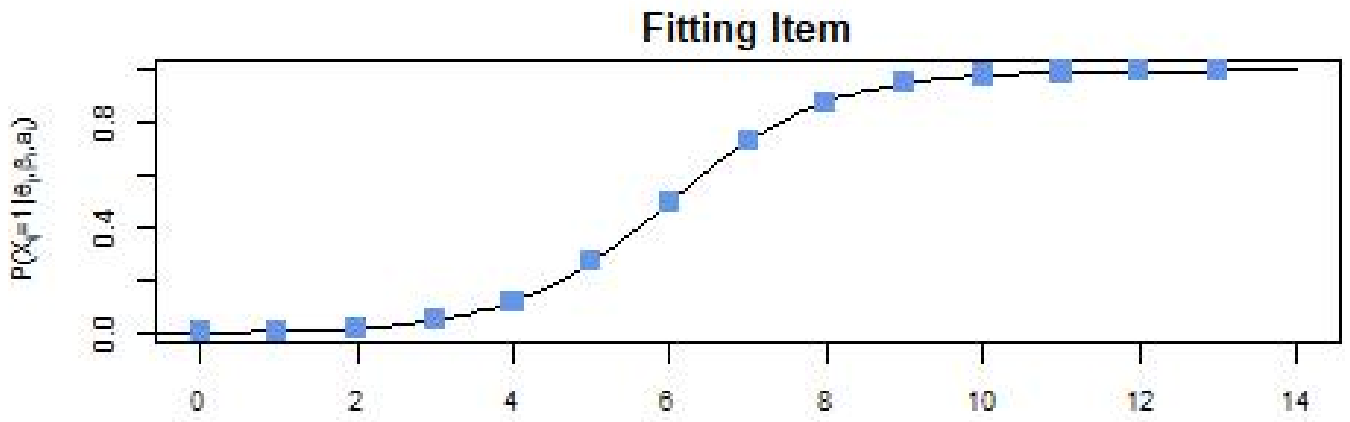
Underfit indicates underdiscrimination, the information is «blurred». It is not possible to differentiate ability levels. Underfit is found when the Infit or Outfit are much above 1.

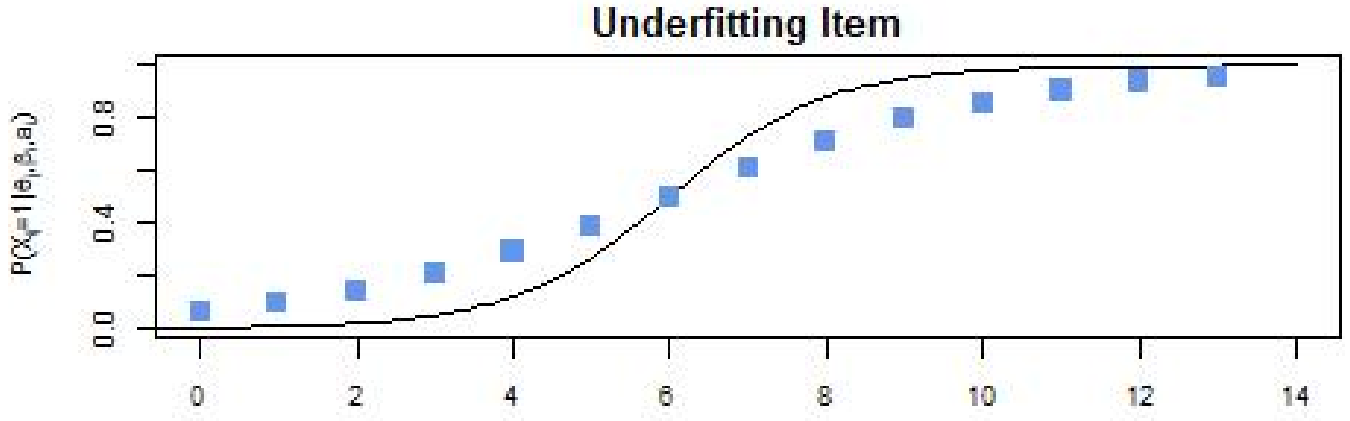
Overfit indicates overderdiscrimination, the information is too «sharp». An overdiscriminating item acts like an on-off switch. Overfit is found when the Infit or Outfit are much below 1.

Note:

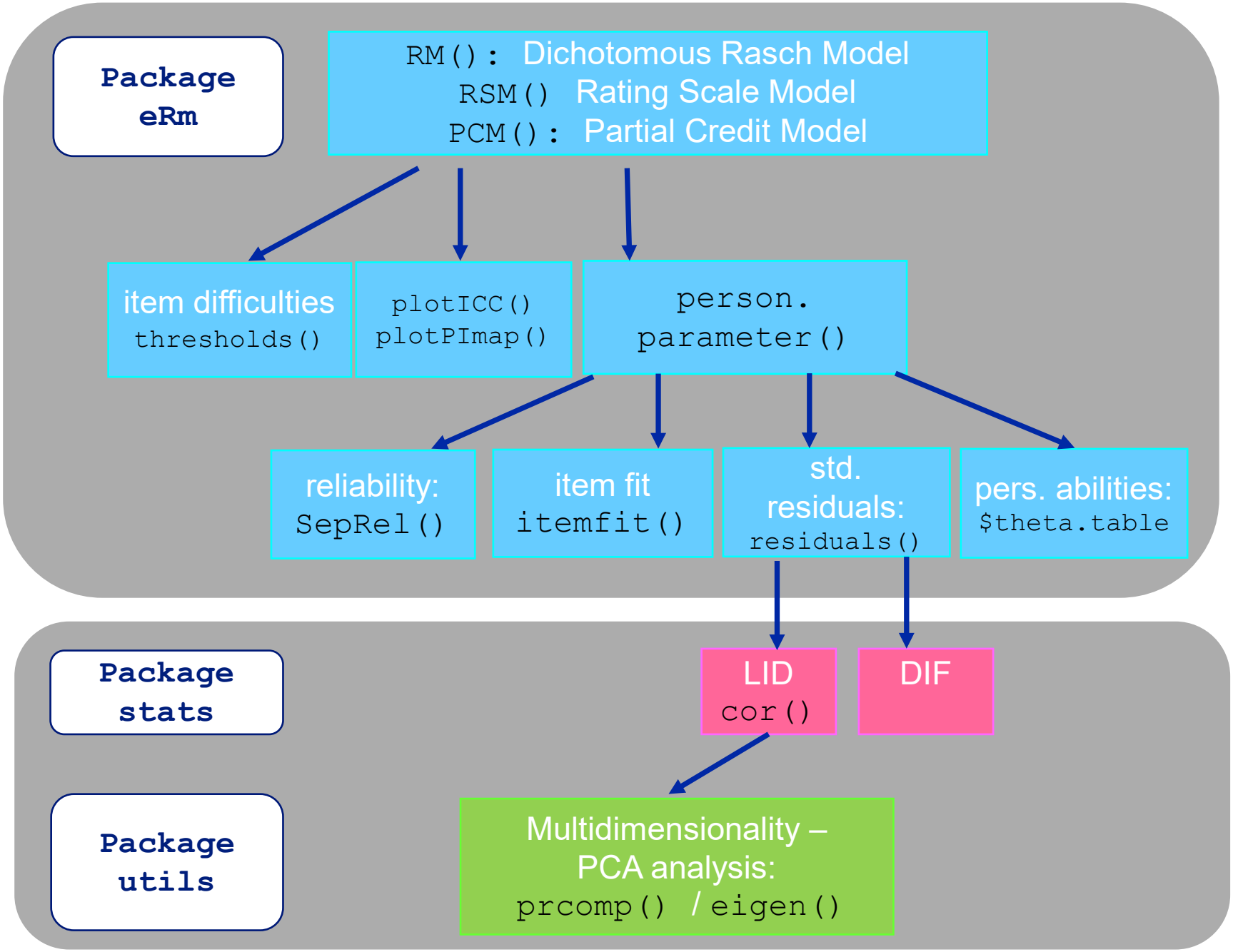
- Overfit is less critical for scales than underfit.

- Cut-off for acceptable fit, in terms of how much underfit can be tolerated, depends on the purpose of a scale.


Fitting Item
P(X_q=1|θ_i,β_i,a_i)
0.0
0.4
0.8
0
2
4
6
8
10
12
14


Underfitting Item
P(X_q=1|θ_i,β_i,a_i)
0.0
0.4
0.8
0
2
4
6
8
10
12
14

Overfitting Item
P(X_q=1|θ_i,β_i,a_i)
0.0
0.4
0.8
0
2
4
6
8
10
12
14


Package eRm
RM(): Dichotomous Rasch Model
RSM() Rating Scale Model
PCM(): Partial Credit Model
item difficulties
thresholds()
plotICC()
plotPImap()
person.
parameter()
reliability:
SepRel()
item fit
itemfit()
std.
residuals:
residuals()
pers. abilities:
$theta.table
Package stats
LID
cor()
DIF
Package utils
Multidimensionality –
PCA analysis:
prcomp() / eigen()

Let’s go to R-Studio

Open the R-Script TT3_Rscript.r that you find in Github

Exercise

Smith (1998, p.78) suggests cut-off based on sample sizes. Have a look at the article. Write down the two formulas and calculate the reasonable cut-offs for infit and outfit. Discuss item fit of the SRG-data under this light.