Package ‘pairwise’

September 29, 2017

Type Package
Maintainer Joerg-Henrik Heine <jhheine@googlemail.com>
Date 2017-09-29
Author Joerg-Henrik Heine <jhheine@googlemail.com>
Version 0.4.3-2
Encoding UTF-8
License GPL-3
Imports graphics, stats, methods
Depends R (>= 2.10.1)
Title Rasch Model Parameters by Pairwise Algorithm
Description Performs the explicit calculation
    -- not estimation! -- of the Rasch item parameters for dichotomous and
    polytomous item responses, using a pairwise comparison approach. Person
    parameters (WLE) are calculated according to Warm’s weighted likelihood
    approach.
Suggests psych
RoxygenNote 6.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2017-09-29 15:26:43 UTC

R topics documented:

pairwise-package .............................................. 3
andersentest.pers ........................................... 5
bfiN .......................................................... 6
bfiN_miss ..................................................... 8
bfi_cov ......................................................... 9
catprob ......................................................... 10
cog ............................................................ 11
cogBOOKLET ................................................... 12
<table>
<thead>
<tr>
<th>R topics documented:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>deltapar</td>
<td>12</td>
</tr>
<tr>
<td>DEU_PISA2012</td>
<td>14</td>
</tr>
<tr>
<td>esc</td>
<td>14</td>
</tr>
<tr>
<td>ftab</td>
<td>15</td>
</tr>
<tr>
<td>gif</td>
<td>16</td>
</tr>
<tr>
<td>grm</td>
<td>17</td>
</tr>
<tr>
<td>iff</td>
<td>20</td>
</tr>
<tr>
<td>kft5</td>
<td>21</td>
</tr>
<tr>
<td>logLik.pers</td>
<td>22</td>
</tr>
<tr>
<td>lrtest.pers</td>
<td>23</td>
</tr>
<tr>
<td>make.incidenz</td>
<td>23</td>
</tr>
<tr>
<td>Neofit5</td>
<td>24</td>
</tr>
<tr>
<td>pair</td>
<td>25</td>
</tr>
<tr>
<td>pairSE</td>
<td>27</td>
</tr>
<tr>
<td>pairwise.item.fit</td>
<td>29</td>
</tr>
<tr>
<td>pairwise.person.fit</td>
<td>30</td>
</tr>
<tr>
<td>pairwise.S</td>
<td>31</td>
</tr>
<tr>
<td>pers</td>
<td>34</td>
</tr>
<tr>
<td>plot.grm</td>
<td>36</td>
</tr>
<tr>
<td>plot.pair</td>
<td>37</td>
</tr>
<tr>
<td>plot.pairSE</td>
<td>37</td>
</tr>
<tr>
<td>plot.pers</td>
<td>38</td>
</tr>
<tr>
<td>plot.rfa</td>
<td>39</td>
</tr>
<tr>
<td>ptbis</td>
<td>40</td>
</tr>
<tr>
<td>Q</td>
<td>41</td>
</tr>
<tr>
<td>q3</td>
<td>42</td>
</tr>
<tr>
<td>rfa</td>
<td>43</td>
</tr>
<tr>
<td>sim200x3</td>
<td>44</td>
</tr>
<tr>
<td>simra</td>
<td>45</td>
</tr>
<tr>
<td>summary.grm</td>
<td>46</td>
</tr>
<tr>
<td>summary.pair</td>
<td>47</td>
</tr>
<tr>
<td>summary.pair.S</td>
<td>47</td>
</tr>
<tr>
<td>summary.pairSE</td>
<td>48</td>
</tr>
<tr>
<td>summary.pairwise.item.fit</td>
<td>48</td>
</tr>
<tr>
<td>summary.pairwise.person.fit</td>
<td>49</td>
</tr>
<tr>
<td>summary.pers</td>
<td>49</td>
</tr>
<tr>
<td>summary.rfa</td>
<td>50</td>
</tr>
<tr>
<td>tff</td>
<td>50</td>
</tr>
</tbody>
</table>

Index 52
Description

Performs the explicit calculation – not estimation! – of the Rasch item parameters for dichotomous and polytomous response formats using a pairwise comparison approach (Choppin, 1968, 1985). On the basis of the item parameters, person parameters (WLE) are calculated according to Warm’s weighted likelihood approach (Warm, 1989). Item- and person fit statistics and several functions for plotting are available.

Details

In case of dichotomous answer formats the item parameter calculation for the Rasch Model (Rasch, 1960), is based on the construction of a pairwise comparison matrix $M_{ij}$ with entries $f_{ij}$ representing the number of respondents who got item $i$ right and item $j$ wrong according to Choppin’s (1968, 1985) conditional pairwise algorithm.

For the calculation of the item thresholds and difficulty in case of polytomous answer formats, according to the Partial Credit Model (Masters, 1982), a generalization of the pairwise comparison algorithm is used. The construction of the pairwise comparison matrix is therefore extended to the comparison of answer frequencies for each category of each item. In this case, the pairwise comparison matrix $M_{nicjc}$ with entries $f_{icjc}$ represents the number of respondents who answered to item $i$ in category $c$ and to item $j$ in category $c-1$ widening Choppin’s (1968, 1985) conditional pairwise algorithm to polytomous item response formats. Within R this algorithm is simply realized by matrix multiplication.

In general, for both polytomous and dichotomous response formats, the benefit in applying this algorithm lies in it’s capability to return stable item parameter 'estimates' even when using data with a relative high amount of missing values, as long as the items are still proper linked together.

The recent version of the package ‘pairwise’ computes item parameters for dichotomous and polytomous item responses – and a mixture of both – according the partial credit model using the function `pair`.

Based on the explicit calculated item parameters for a dataset, the person parameters may thereupon be estimated using any estimation approach. The function `pers` implemented in the package uses Warm’s weighted likelihood approach (WLE) for estimation of the person parameters (Warm, 1989). When assessing person characteristics (abilities) using (rotated) booklet designs an ‘incidence’ matrix should be used, giving the information if the respective item was in the booklet (coded 1) given to the person or not (coded 0). Such a matrix can be constructed (out of a booklet allocation table) using the function `make.incidenz`.

Item- and person fit statistics, see functions `pairwise.item.fit` and `pairwise.person.fit` respectively, are calculated based on the squared and standardized residuals of observed and the expected person-item matrix. The implemented procedures for calculating the fit indices are based on the formulas given in Wright & Masters, (1982, p. 100), with further clarification given at http://www.rasch.org/rmt/rmt34e.htm.

Further investigation of item fit can be done by using the function `ptbis` for point biserial correlations. For a graphical representation of the item fit, the function `gif` for plotting empirical and
model derived category probability curves, or the function `esc` for plotting expected (and empirical) score curves, can be used.

The function `iff` plots or returns values of the item information function and the function `tff` plots or returns values of the test information function.

To detect multidimensionality within a set of Items a rasch residual factor analysis proposed by Wright (1996) and further discussed by Linacre (1998) can be performed using the function `rfa`. For a 'heuristic' model check the function `grm` makes the basic calculations for the graphical model check for dicho- or polytomous item response formats. The corresponding S3 plotting method is `plot.grm`.

**Author(s)**

Joerg-Henrik Heine <jhheine@googlemail.com>

**References**


Andersen's Likelihood Ratio Test for Object of class "pers"

Description

The Andersen likelihood ratio test is based on splitting the dataset into subgroups of persons. One can argue that it is a significance testable version of the more descriptive graphical model check - see grm.

Usage

`andersentest.pers(pers_obj, split = "median", splitseed = "no", pot = NULL, zerocor = NULL)`

Arguments

- **pers_obj**: an object of class "pers" - see function `pers`.
- **split**: Specifies the splitting criterion. Basically there are three different options available - each with several modes - which are controlled by passing the corresponding character expression to the argument.
  1) Using the rawscore for splitting into subsamples with the following modes:
     - `split = "median"`: median raw score split - high score group and low score group;
     - `split = "mean"`: mean raw score split - high score group and low score group. Finally `split = "score"` that is splitting `daten` into as many subsamples as there are raw score groups - discarding min and max (theoretical) score group - which matches the concept proposed by Andersen (1973).
  2) Dividing the persons in `daten` into subsamples with equal size by random allocation with the following modes:
     - `split = "random"` (which is equivalent to `split = "random.2"`) divides persons into two subsamples with equal size. In general the number of desired subsamples must be expressed after the dot in the character expression - e.g. `split = "random.6"` divides persons into 6 subsamples (with equal size) by random allocation etc.
  3) The third option is using a manifest variable as a splitting criterion. In this case a vector with the same length as number of cases in `daten` must be passed to the argument grouping the data into subsamples. This vector should be coded as "factor" or a "numeric" integer vector with min = 1.
- **splitseed**: numeric, used for `set.seed(splitseed)` for random splitting - see argument `split`.
- **pot**: optional argument, at default (`pot=NULL`) setting is read from `pers_obj` - see description for `pair`.
- **zerocor**: optional argument, at default (`zerocor=NULL`) setting is read from `pers_obj` - see description for `pair`.

more...
Details

Andersen (1973) proposed to split the dataset by [raw] score groups, which can be achieved setting the argument `split = "score"`. However, as pointed out by Rost (2004) there might be several different splitting criteria for testing subsample invariance of the Rasch model. Thus the argument `split` provides some other options for splitting the data - see description of arguments.

Value

A (list) object of class "andersentest.pers" ...

References


Examples

```r
## Not run:
data(bfiN) # loading example data set
data(bfi_cov) # loading covariates to bfiN data set
model <- pers(pair(bfiN,m=6))
andersentest.pers(model, split = bfi_cov$gender)
andersentest.pers(model, split = "random")
andersentest.pers(model, split = "median")
## Unsing simulated data:
data("sim200x3")
model2 <- pers(pair(sim200x3))
andersentest.pers(model2, split = "median")
## End(Not run)
```

---

**bfiN**

*5 polytomous personality items*

Description

Data from 2800 subjects answering to 5 neuroticism items with 6 answer categories (0-5) of the *bfi* dataset originally included in the R-package `{psych}` - see [https://cran.r-project.org/package=psych](https://cran.r-project.org/package=psych).

Usage

data(bfiN)

Format

A "data.frame" containing 5 variables and 2800 observations.
Details

The other variables from the original bfi dataset were skipped and the categories are 'downcoded' to '0,1,2,3,4,5' to have a simple, ready to use example data frame. For further Information on the original dataset see R-package {psych}.

The category meanings (after downcoding) are as follows:

score 0 Very Inaccurate
score 1 Moderately Inaccurate
score 2 Slightly Inaccurate
score 3 Slightly Accurate
score 4 Moderately Accurate
score 5 Very Accurate

The Item meanings are as follows:

N1 Get angry easily.
N2 Get irritated easily.
N3 Have frequent mood swings.
N4 Often feel blue.
N5 Panic easily.

The covariates like gender, education and age are in a separate dataset cov_bfi

Source

https://cran.r-project.org/package=psych

References

Revelle, William (2015), psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 1.5.1

Examples

data(bfiN)
dim(bfiN)

names(bfiN) # show all variable names of data.frame bfiN
range(bfiN,na.rm=TRUE) # checking the valid response range
bfiN_miss
5 polytomous personality items

Description
Data from 2800 subjects answering to 5 neuroticism items with 6 answer categories (0-5) of the bfi dataset originally included in the R-package {psych} with artificial missing data (see details).

Usage
data(bfiN_miss)

Format
A "data.frame" containing 5 variables and 2800 observations.

Details
This dataset is the same like the dataset {bfiN} included in this package, except for the amount of missing data, which were additional created in that way, having approx. 15% missing for each of the 5 variables by random.
The other variables from the original bfi dataset were skipped and the categories are 'downcoded' to '0,1,2,3,4,5' to have a simple, ready to use example data frame. For further Information on the original dataset see R-package {psych}. The covariates like gender, education and age are in a separate dataset cov_bfi

Source
https://cran.r-project.org/package=psych

References
Revelle, William (2015), psych: Procedures for Psychological, Psychometric, and Personality Research.R package version 1.5.1

Examples
data(bfiN_miss)
dim(bfiN_miss)
# show all variable names of data.frame bfiN_miss
names(bfiN_miss)
# checking the valid response range
range(bfiN_miss, na.rm=TRUE)
# percentage of missing per variable
colSums(is.na(bfiN_miss))/dim(bfiN_miss)[1]
Covariates to the data from 2800 subjects answering to 5 neuroticism items of the bfi dataset originally included in the R-package \{} \texttt{psych} \texttt{\}} - see \url{https://cran.r-project.org/package=psych}.

Usage

\texttt{data(bfi_cov)}

Format

A \texttt{data.frame} containing 3 variables (gender, education, and age) for 2800 observations.

Details

The covariates are in the same row (person) order as the responses to the 5 neuroticism items in the separate datasets \texttt{bfiN} and \texttt{bfiN_miss}. The coding is as follows:

- \texttt{gender} Males = 1, Females = 2
- \texttt{education} 1 = HS, 2 = finished HS, 3 = some college, 4 = college graduate 5 = graduate degree
- \texttt{age} age in years

Source

\url{https://cran.r-project.org/package=psych}

References

Revelle, William (2015), psych: Procedures for Psychological, Psychometric, and Personality Research. \textit{R package version 1.5.1}

Examples

\texttt{data(bfi_cov)}
\texttt{dim(bfi_cov)}
\texttt{# show all variable names of data}
**Description**

plotting function for plotting category probability curves.

**Usage**

catprob(pair_obj, itemnumber = 1, ra = 4, plot = TRUE, ...)

**Arguments**

- **pair_obj**: an object of class "pair" as a result from function `pair`.
- **itemnumber**: an integer, defining the number of the item to plot the respective category probability for. This is set to an arbitrary default value of `itemnumber = 1` to avoid error messages when you forget to choose an item to plot the expected score curves for.
- **ra**: an integer, defining the (logit) range for x-axis
- **plot**: a logical (default `plot = TRUE`), defining wether to supress plotting an just return a matrix of category probabilities
- **...**: arguments passed to `plot`

**Details**

no details in the moment.

**Value**

a plot or a matrix with category probabilities.

**Examples**

```r
###
data(sim200x3)
result <- pair(sim200x3)
catprob(pair_obj = result, itemnumber = 2)
data(bfiN)
result <- pair(bfiN)
catprob(pair_obj = result, itemnumber = 3)
```
Math PISA (2003) data

Description

Data from the german sample of the PISA 2003 survey, containing 31 dichotomous items from the math task.

Usage

data(cog)

Format

A data frame containing 34 variables and 4660 observations.

Details

The first 3 variables are ID variables. For further Information on variables and their meaning see the codebook PDF file available at https://www.oecd.org/pisa/data/

Source

https://www.oecd.org/pisa/data/

References


Examples

data(cog)
dim(cog)

names(cog) # show all variable names of data.frame cog
names(cog[,4:34]) # show the variable names of the math items
names(cog[,1:3]) # show the variable names of the ID variables
**cogBOOKLET**

*Booklet allocation table for Math PISA (2003) data*

**Description**

A `data.frame` containing a booklet allocation table for the cognitive Data `cog` in this package, which holds 31 dichotomous items from the math task from the german sample of the PISA 2003 survey.

**Usage**

```r
data(cogBOOKLET)
```

**Format**

A `data.frame` containing 31 rows.

**Details**

For further information on variables and their meaning see the codebook PDF file available at [https://www.oecd.org/pisa/data/](https://www.oecd.org/pisa/data/)

**Source**

[https://www.oecd.org/pisa/data/](https://www.oecd.org/pisa/data/)

**References**


**Examples**

```r
data(cogBOOKLET)
cogBOOKLET
```

deltapar

*Compute delta parameters from thurstonian thresholds*

**Description**

Calculation of delta parameters or rather item step parameters from thurstonian threshold parameters returned by the function `pair`.

**Usage**

```r
deltapar(object, sigma = TRUE)
```
**deltapar**

**Arguments**

- **object**: an object of class "pair" as resulting from item parameter calculation using the function `pair`
- **sigma**: a logical whether to return item difficulties (sigma) or not

**Details**

The "Thurstone threshold" or rather thurstonian threshold for a category corresponds to a point on the latent variable at which the probability of being observed in that category or above equals that of being observed in the categories below. Thus these thurstonian threshold parameters can be interpreted in an strait forward and easy way. However, some other computer programs related to Rasch analysis don’t return thurstonian threshold parameters from their estimation procedure, but rather so called delta parameters for the item steps. The later are also known as "step measures", "step calibrations", "step difficulties", "tau parameters", and "Rasch-Andrich thresholds". For a better comparability between different Rasch software and estimation procedures the thurstonian threshold parameters can be converted into delta or rather items step parameters.

**Value**

If `sigma=TRUE` an object of class c("data.frame", "deltapar") containing delta parameters for items and their difficultie (first column). Otherwise a matrix containing only the delta parameters.

**References**


**Examples**

```
# Loading response data
data(sim200x3)  # loading reponse data
ip <- pair(sim200x3, m = c(2,3,3))  # compute item parameters
summary(ip)  # have a look at the results (thurstonian thresholds)
deltapar(ip)  # compute delta parameters from these
```
DEU_PISA2012  Data from PISA 2012 - German Sample

Description
Selected data for 5001 'subjects' who participated in the PISA 2012 survey.

Usage
data(DEU_PISA2012)

Format
A list containing ...

Details
The data is based on freely downloadable data on the official OECD page - see source. The general structure of the data in list format, is described in an PDF document available in the User guides, package vignettes and other documentation section.

Source
https://www.oecd.org/pisa/data/

References
To come ...

Examples

# Expected Score Curves Plots

data(DEU_PISA2012)
str(DEU_PISA2012)

Description
plotting function for plotting expected score curves.

Usage
esc(pers_obj, itemnumber = 1, integ = 6, ra = 4, nodes = 100, lwd = 2, ...)

Arguments

pers_obj an object of class "pers" as a result from function pers.

itemnumber an integer, defining the number of the item to plot the respective category probability for. This is set to an arbitrary default value of itemnumber = 1 to avoid error messages when you forget to choose an item to plot the expected score curves for.

integ either an integer defining the number of (ability) groups to integrate the empirical theta vector or the character expression "all" to plot the empirical theta distribution at the respective item score using symbols (see example).

ra an integer, defining the (logit) range for x-axis

nodes number of integration nodes

lwd see plot

... arguments passed to plot

Details

no details in the moment.

Examples

#########
data(bfiN)
result <- pers(pair(bfiN))
esc(pers_obj=result,1,lwd=2) # plot for first item
esc(pers_obj=result,2,lwd=2) # plot for second item
for(i in 1:5){esc(pers_obj=result,i,lwd=2)}
#########
esc(pers_obj=result,2,integ="all",lwd=2) # plot for second item

ftab Tabulating Answer Categories in Data

Description

function tabulating (answer) categories in X.

Usage

ftab(X, categories = NULL, na.omit = FALSE)
Arguments

- **X**: Data as a "matrix", a "data.frame" or even a "vector" or "factor". "vector" or "factor" are coerced to a "data.frame" with one column.
- **categories**: Optional a vector ("numeric" or "character") containing the categories to tabulate. At default (categories=NULL) the function looks for unique categories in X.
- **na.omit**: Logical (default: na.omit=FALSE) whether to return frequencies for missing values, NAs.

Details

X can either be a ("numeric" or "character") "matrix" containing response vectors of persons (rows) or a "data.frame" containing "numeric", "character" or "factor" variables (columns).

Value

- A "matrix" with category frequencies

Examples

```r
# Example usage
data(bfiN)
ftab(bfiN)
data(sim200x3)
ftab(sim200x3)
```

### gif

**Graphical Item Fit Plots**

**Description**

Plotting function for plotting empirical and model derived category probability curves.

**Usage**

```
gif(pers_obj, itemnumber = 1, ra = 4, integ = "raw", kat = "all", ...)
```

**Arguments**

- **pers_obj**: An object of class "pers" as a result from function `pers`.
- **itemnumber**: An integer, defining the number of the item to plot the respective category probability for. This is set to an arbitrary default value of `itemnumber = 1` to avoid error messages when you forget to choose an item to plot the expected score curves for.
- **ra**: An integer, defining the (logit) range for x-axis.
either an integer, defining the number of integration points along the (logit) range on the x-axis to integrate the empirical theta values, or the character expression "raw" (default) which will use the rawscore groups as integration points.

either an integer, defining for which category the empirical category probabilities should be plotted over the model derived category probability curves, or the character expression "all" (default) which will plot the empirical category probabilities for all categories.

... arguments passed to plot

Details

no details in the moment.

Value

a plot with category probabilities.

Examples

```r
###
data(bfiN)
pers_obj <- pers(pair(bfiN))
### plot empirical category probabilities
gif(pers_obj = pers_obj, itemnumber = 1 )
gif(pers_obj = pers_obj, itemnumber = 1 , integ=8) # integration over 8 points
gif(pers_obj = pers_obj, itemnumber = 1 , integ=8, kat=1) # only for category number 1
```

grm

Graphical Model Check

Description

This function makes the basic calculations for the graphical model check for dicho- or polytomous item response formats. It is more or less a wrapper function, internaly calling the function pairSE. Several splitting options are available (see arguments).

Usage

```r
grm(daten, m = NULL, split = "random", splitseed = "no",
     verbose = FALSE, ...)
```
Arguments

daten

A data frame or matrix with optionally named columns (names of items), potentially with missing values, comprising polytomous or dichotomous (or mixed category numbers) responses of \( n \) respondents (rows) on \( k \) items (columns) coded starting with 0 for lowest category to \( m-1 \) for highest category, with \( m \) being a vector (with length \( k \)) with the number of categories for the respective item.

\( m \)

An integer (will be recycled to a vector of length \( k \)) or a vector giving the number of response categories for all items - by default \( m = \) \text{NULL}, \( m \) is calculated from data, assuming that every response category is at least once present in data. For sparse data it is strongly recommended to explicitly define the number of categories by defining this argument.

split

Specifies the splitting criterion. Basically there are three different options available - each with several modes - which are controlled by passing the corresponding character expression to the argument.

1) Using the rawscore for splitting into subsamples with the following modes:

- \( \text{split} = \) "median" median raw score split - high score group and low score group;
- \( \text{split} = \) "mean" mean raw score split - high score group and low score group.

2) Dividing the persons in \( \text{daten} \) into subsamples with equal size by random allocation with the following modes:

- \( \text{split} = \) "random" (which is equivalent to \( \text{split} = \) "random.2") divides persons into two subsamples with equal size. In general the number of desired subsamples must be expressed after the dot in the character expression - e.g. \( \text{split} = \) "random.6" divides persons into 6 subsamples (with equal size) by random allocation etc.

3) The third option is using a manifest variable as a splitting criterion. In this case a vector with the same length as number of cases in \( \text{daten} \) must be passed to the argument grouping the data into subsamples. This vector should be coded as "factor" or a "numeric" integer vector with min = 1.

splitseed

Numeric, used for \text{set.seed(splitseed)} for random splitting - see argument \text{split}.

verbose

Logical, if \text{verbose} = \text{TRUE} (default) a message about subsampling is sent to console when calculating standard errors.

Details

The data is splitted in two or more subsamples and then item thresholds, the parameter (Sigma) and their standard errors (SE) for the items according the PCM are calculated for each subsample. Additional arguments (see description of function \text{pairSE}) for parameter calculation are passed through.

WARNING: When using data based on booklet designs with systematically missing values (by design) you have to ensure that in each of the booklet the maximum raw value to reach is equal while using the raw value as splitting criterion.
Value

A (list) object of class "grm" containing the item difficulty parameter sigma and their standard errors for two or more subsamples.

A note on standard errors

Estimation of standard errors is done by repeated calculation of item parameters for subsamples of the given data. This procedure is mainly controlled by the arguments nsample and size (see arguments). With regard to calculation time, the argument nsample is the 'time killer'. On the other hand, things (estimation of standard errors) will not necessarily get better when choosing large values for nsample. For example choosing nsample=400 will only result in minimal change for standard error estimation in comparison to (nsample=30) which is the default setting (see examples).

References

description of function pairSE{pairwise}.

Examples

data(bfiN) # loading example data set
data(bfi_cov) # loading covariates to bfiN data set

# calculating itemparameters and SE for two random allocated subsamples
grm_gen <- grm(daten=bfiN, split = bfi_cov$gender)
summary(grm_gen)
plot(grm_gen)

grm_med <- grm(daten=bfiN, split = "median")
summary(grm_med)
plot(grm_med)

grm_ran<grm(daten=bfiN, split = "random")
summary(grm_ran)

# some examples for plotting options
# plotting item difficulties for two subsamples against each other
# with ellipses for a CI = 95%.
plot(grm_ran)

# using triangles as plotting pattern
plot(grm_ran,pch=2)

#plotting without CI ellipses
plot(grm_ran,ci=0,pch=2)

# plotting with item names
plot(grm_ran,itemNames=TRUE)
# Changing the size of the item names
plot(grm_ran, itemNames=TRUE, cex.names = 1.3)

# Changing the color of the CI ellipses
plot(grm_ran, itemNames=TRUE, cex.names = .8, col.error="green")

AAAA### example from details section 'Some Notes on Standard Errors'AAAA###
grm_def<-grm(daten=bfiN, split = "random", splitseed=13)
plot(grm_def)

AAAA###
grm_400<-grm(daten=bfiN, split = "random", splitseed=13 ,nsample=400)
plot(grm_400)


tttt

iff

Item information function

Description
plotting function for plotting the Item information function(IIF).

Usage

iff(pair_obj, itemnumber = 1, x = NULL, plot = TRUE, cat = FALSE, 
lwd = 2, col = 1, ...)

Arguments

pair_obj an object of class "pair" as a result from function pair.
itemnumber an integer, defining the number of the item to plot the respective item information function for. This is set to an arbitrary default value of itemnumber = 1 to avoid error messages when you forget to choose an item to plot the item information function for.
x The value(s) of the latent variable, at which the IIF will be evaluated. x should be either a numeric vector of theta values or a single numeric value. If x is given as a single numeric value plotting is supressed. If not given (default), 99 values spaced evenly between -4 and +4 will be used, handy for plotting.
plot a logical (default plot = TRUE), defining wether to supress plotting an just return a matrix of the values of the Item information function.
cat a logical (default cat = FALSE), defining wether to plot or return the values of the Item information function based on item categories.
lwd see parameters for plot
col see parameters for plot
...
arguments passed to plot
Details

no details in the moment.

Value

a plot, a matrix or a single numeric with values of the Item information function.

Examples

```
####
data(sim200x3)
result <- pair(sim200x3)
# IFF plot for Item No. 2
iff(pair_obj = result, itemnumber = 2)
# IFF plot for Categories of Item No. 2
iff(pair_obj = result, itemnumber = 2, cat=TRUE)
# IFF at theta=0 for Item No. 2
iff(pair_obj = result, itemnumber = 2, x=0)
# IFF at theta=0 for Categories of Item No. 2
iff(pair_obj = result, itemnumber = 2, x=0, cat=TRUE)
# IFF of Item No. 2 for a given range of thetas
iff(pair_obj = result, itemnumber = 2, x=seq(0,4,.1))
# ... etc.
iff(pair_obj = result, itemnumber = 2, x=seq(0,4,.1), cat=TRUE)
#### examples with other data ...
data(bfiN)
result <- pair(bfiN)
iff(pair_obj = result, itemnumber = 3)
iff(pair_obj = result, itemnumber = 3, cat=TRUE)
```

**kft5**  
**Dichotomous example data in Rost 2004**

Description

Data for 300 subjects answering to 5 dichotomous items out of 'Kognitiver Fähigkeits Test' [Cognitive Skills Test] (KFT - Gaedike & Weinläder, 1976). This data is used as an example in the textbook by J. Rost (2004) to demonstrate some principles of rasch measurement.

Usage

data(kft5)

Format

A “matrix” containing 5 columns (variables) and 300 rows (observations).
Details

The instrument KFT and the data are described in Rost (2004) at page 95.

References


Examples

data(kft5)
dim(kft5)

# frequencies
ftab(kft5)

# Item parameter to be compared with Rost (2004), page 120.
summary(pair(kft5))

# Item parameter to be compared with Rost (2004), page 120.
summary(pers(pair(kft5)))

---

**logLik.pers**

*S3 logLik for Object of class "pers"*

Description

S3 logLik method to extract the log-likelihood for object of class "pers"

Usage

```
## S3 method for class 'pers'
logLik(object, sat = FALSE, p = FALSE, ...)
```

Arguments

- **object**
  - object of class "pers"

- **sat**
  - a "logical" with default set to sat=FALSE to return the Log-Likelihood of the data for the unrestricted model based on parameters estimated with function pers. If set to sat=TRUE the Log-Likelihood of the saturated model is returned instead.

- **p**
  - a "logical" with default set to p=FALSE to return the category propabilities for the empirical data.

... not used jet.
lrtest.pers

Likelihood Ratio Test for Object of class "pers"

Description
Function to perform a likelihood ratio test for the estimated model 'against' the saturated model for object of class "pers".

Usage
lrtest.pers(object, ...)

Arguments
object an object of class "pers" - see function pers.
... not used jet.

make.incidenz
Converting a booklet allocation table into a incidence matrix

Description
This function converts a booklet allocation table (like in cogBOOKLET) into a incidence matrix used in the function pers.

Usage
make.incidenz(tab, bookid, item_order = NULL, info = FALSE)

Arguments
tab a booklet allocation table as a data.frame. The first column is assumed to contain the item names as a character vector (not a factor!) the other columns must be integer vectors containing the information in which booklet(s) the respective item is allocated.
bookid a integer vector with the same length as the number of persons in the response data giving the information which booklet was assigned to each person.
item_order optional a character vector with the item names in the order of the items in the response data (from first to last column in the response data). By default it is assumend that the item order in the booklet allocation table is already the same as in the response data.
info logical default: info=FALSE to return just the incidence matrix. If set to info=TRUE more detailed information about the booklet design ist returned.
Details

It is assumed that there is an equal replicate factor for each item used, when constructing the booklet design - so every item occurs with the same frequency over all booklets of the entire set of booklets.

Value

an incidence matrix as an object of class "matrix" with 0,1 coding or a "list" with detailed information.

Examples

#################################
data(cog);data(cogBOOKLET) # loading response and allocation data
table(cog$BOOKID)# show n persons per booklet
names(table(c(as.matrix(cogBOOKLET[,2:5])))) # show booklets in allocation data
d<-c(cog[cog$BOOKID!=14,]) # skip persons which got booklet No.14.
inc<-make.incidenz(tab=cogBOOKLET, bookid=d$BOOKID) # make just the incidence matrix
inc
make.incidenz(tab=cogBOOKLET, bookid=d$BOOKID, info=TRUE) # get some info too
# in this case not necessary but just to show
# using the (item) names in cog to secure the item order in incidence matrix:
make.incidenz(tab=cogBOOKLET, bookid=d$BOOKID, item_order=names(cog)[4:34])
#################################

Neoffi5

Polytomous example data in Rost 2004

Description

Data for 1000 subjects answering to 5 polytomous items assessing neuroticism contained in the german version of the NEO–five–factor–inventory (NEOFFI) by Borkenau and Ostendorf (1991). This data is used as an example in the textbook by J. Rost (2004) to demonstrate some principles of rasch measurement.

Usage

data(Neoffi5)

Format

A "matrix" containing 5 columns (variables) and 1000 rows (observations).

Details

An detailed description of the data can be found in Rost (2004) at page 202.
pair

References


Examples

data(Neoffi5)
dim(Neoffi5)

# frequencies
ftab(Neoffi5)

# Itemparameter to be compared with Rost (2004), page 211.
summary(pair(Neoffi5))

# Itemparameter to be compared with Rost (2004), page 213.
summary(pers(pair(Neoffi5)))

---

**pair**

Rasch Item Parameter (Main Function)

Description

This is the (new) main function for calculation of the item parameter for the dichotomous Rasch Model (Rasch, 1960) and its extension for polytomous items (thurstonian thresholds) according to the Partial Credit Model (Masters, 1982), using a generalization of the pairwise comparison algorithm (Choppin, 1968, 1985; Wright & Masters, 1982). The number of (response) categories may vary across items. Missing values up to an high amount in data are allowed, as long as items are proper linked together.

Usage

```r
pair(daten, m = NULL, pot = TRUE, zerocor = TRUE, ccf = FALSE, ...)
```

Arguments

daten

A data frame or matrix with optionally named columns (names of items), potentially with missing values, comprising polytomous or dichotomous (or mixed category numbers) responses of n respondents (rows) on k items (columns) coded starting with 0 for lowest category to m-1 for highest category, with m being a vector (with length k) with the number of categories for the respective item.

m

An integer (will be recycled to a vector of length k) or a vector giving the number of response categories for all items - by default \( m = \text{NULL} \), m is calculated from data, assuming that every response category is at least once present in data. For 'sparse' data it is strongly recommended to explicitly define the number of categories by defining this argument.
either a logical or an integer >= 2 defining the power to compute of the pairwise comparison matrix. If TRUE (default) a power of three of the pairwise comparison matrix is used for further calculations. If FALSE no powers are computed.

zerocor logical, if TRUE (default) unobserved combinations (1-0, 0-1) in data for each pair of items are given a frequency of one conf. proposal by Alexandrowicz (2011, p.373).

ccf logical with default ccf=FALSE to perform normal item parameter calculation, if set to ccf=TRUE just the conditional item (category) frequencies are returned.

... additional parameters passed through.

Details

Parameter calculation is based on the construction of a paired comparison matrix $M_{nicjc}$ with entries $f_{icjc}$ representing the number of respondents who answered to item $i$ in category $c$ and to item $j$ in category $c-1$ widening Choppin’s (1968, 1985) conditional pairwise algorithm to polytomous item response formats. This algorithm is simply realized by matrix multiplication.

To avoid numerical problems with off diagonal zero’s when constructing the pairwise comparison matrix $M_{nij}$, powers of the $M_{nicjc}$ matrix, can be used (Choppin, 1968, 1985). Using powers $k$ of $M_{nicjc}$ - argument pot=TRUE (default), replaces the results of the direct comparisons between $i$ and $j$ with the sum of the indirect comparisons of $i$ and $j$ through an intermediate $k$.

In general, it is recommended to use the argument with default value pot=TRUE.

For a graphic representation of the item 'estimates' the plotting S3 method plot.pair is available. For plotting the item category probabilities the function catprob can be used.

Value

A (list) object of class "pair" containing the item category thresholds and difficulties sigma, also called item location.

References


Examples

data(bfiN) # loading example data set
# calculating item parameters for 5 neuroticism items with 6 answer categories (0-5).
neuro_itempar<-pair(daten = bfiN, m = 6)
pairSE

Item Parameter calculation with Standard Errors for polytomous Partial Credit Model

Description

Calculation of the item parameters for dichotomous (difficulty) or polytomous items (thurstonian thresholds) and their standard errors (SE) respectively. All parameters are calculated using a generalization of the pairwise comparison algorithm (Choppin, 1968, 1985). Missing values up to an high amount in data matrix are allowed, as long as items are proper linked together.

Usage

pairSE(daten, m = NULL, nsample = 30, size = 0.5, seed = "no", pot = TRUE, zerocor = TRUE, verbose = TRUE, ...)

Arguments

daten a data.frame or matrix with optionaly named columns (names of items), potentially with missing values, comprising polytomous or dichotomous (or mixed category numbers) responses of n respondents (rows) on k items (columns) coded starting with 0 for lowest category to m-1 for highest category, with m beeing a vector (with length k) with the number of categories for the respective item.

m an integer (will be recycled to a vector of length k) or a vector giving the number of response categories for all items - by default m = NULL, m is calculated from data, assuming that every response category is at least once present in data. For sparse data it is strongly recomended to explicitly define the number of categories by defining this argument.

nsample numeric specifying the number of subsamples sampled from data, which is the number of replications of the parameter calculation.

WARNING! specifying high values for nsample (> 100) may result in long computing time without leading to "better" estimates for SE. This may also be the case when choosing argument size="jack" (see argument size) in combination with large datasets (N > 5000).
size numeric with valid range between 0 and 1 (but not exactly 0 or 1) specifying the size of the subsample of data when bootstrapping for SE estimation. As an alternative, size can be set to the character "jack" (size="jack"). This will set the subsample size to N-1 and set nsample=N (see argument nsample), with N being the number of persons in data.

seed numeric used for set.seed(seed).

pot logical, if TRUE (default) a power of three of the pairwise comparison matrix is used for further calculations.

zerocor logical, if TRUE (default) unobserved combinations (1-0, 0-1) in data for each pair of items are given a frequency of one conf. proposal by Alexandrowicz(2011, p.373).

verbose logical, if verbose = TRUE (default) a message about subsampling is sent to console when calculating standard errors.

... additional parameters passed through.

Details

Parameter calculation is based on the construction of a paired comparison matrix \( M_{nicjc} \) with entries \( f_{icjc} \), representing the number of respondents who answered to item \( i \) in category \( c \) and to item \( j \) in category \( c-1 \) widening Choppin’s (1968, 1985) conditional pairwise algorithm to polytomous item response formats. This algorithm is simply realized by matrix multiplication.

Estimation of standard errors is done by repeated calculation of item parameters for subsamples of the given data.

To avoid numerical problems with off diagonal zeros when constructing the pairwise comparison matrix \( M_{nicjc} \), powers of the \( M_{nicjc} \) matrix, can be used (Choppin, 1968, 1985). Using powers \( k \) of \( M_{nicjc} \), argument pot=TRUE (default), replaces the results of the direct comparisons between \( i \) and \( j \) with the sum of the indirect comparisons of \( i \) and \( j \) through an intermediate \( k \).

In general, it is recommended to use the argument with default value pot=TRUE.

Value

A (list) object of class "pairSE" containing the item category thresholds, difficulties sigma and their standard errors.

A note on standard errors

Estimation of standard errors is done by repeated calculation of item parameters for subsamples of the given data. This procedure is mainly controlled by the arguments nsample and size (see arguments). With regard to calculation time, the argument nsample may be the 'time killer'. On the other hand, things (estimation of standard errors) will not necessarily get better when choosing large values for nsample. For example choosing nsample=400 will only result in minimal change for standard error estimation in comparison to (nsample=30) which is the default setting (see examples).
pairwise.item.fit

References


Examples

```
data(bfiN) # loading example data set

# calculating item parameters and their SE for 5 neuroticism items with 6 answer categories (0-5).
neuro_itempar <- pairwise.itemfit(obj = bfiN, m = 6)
summary(neuro_itempar) # summary for result

# plotting item thresholds with with their CI = 95%
plot(neuro_itempar)
plot(neuro_itempar, sortdif = TRUE)
```

```
neuro_itempar_TPP <- pairwise.itemfit(obj = bfiN, m = 6, sample = TRUE)
plot(neuro_itempar)
plot(neuro_itempar_TPP)
```

```
neuro_itempar_TPP <- pairwise.itemfit(obj = bfiN, m = 6, sortdif = TRUE)
plot(neuro_itempar)
plot(neuro_itempar_TPP)
```

Remarks

```
neuro_itempar_TPP <- pairwise.itemfit(obj = bfiN, m = 6, sample = TRUE)
plot(neuro_itempar)
plot(neuro_itempar_TPP)
```

Description

function for calculating item fit indices. The procedures for calculating the fit indices are based on the formulas given in Wright & Masters, (1982, P. 100), with further clarification given in [http://www.rasch.org/rmt/rmt142m.htm](http://www.rasch.org/rmt/rmt142m.htm).

Usage

`pairwise.item.fit(pers_obj, na_treat = NA)`

Arguments

- **pers_obj**: an object of class "pers" as a result from function `pers`
- **na_treat**: value to be assigned to residual cells which have missing data in the original response matrix. default is set to `na_treat=NA` to ignore these cells in further calculations. An option is to set these residuals to 0 using `na_treat=0`, which implies that they are imputed as 'fitting data', i.e., zero residuals. This can attenuate contrasts (see. [http://www.rasch.org/rmt/rmt142m.htm](http://www.rasch.org/rmt/rmt142m.htm)).
pairwise.person.fit

Details
contrary to many IRT software using MI based item parameter estimation, pairwise will not exclude persons, showing perfect response vectors (e.g. c(0,0,0) for dataset with three variables), prior to the scaling. Therefore the fit statistics computed with pairwise may deviate somewhat from the fit statistics produced by IRT software using MI based item parameter estimation (e.g. R-package eRm), depending on the amount of persons with perfect response vectors in the data.

Value
an object of class c("pairwise_item_fit", "data.frame") containing item fit indices.

References

Examples

```r
# data
data(sim200x3)
result <- pers(pair(sin200x3))
pairwise.item.fit(pers_obj=result) # item fit statistic
```

Description
function for calculating person fit indices. The procedures for calculating the fit indices are based on the formulas given in Wright & Masters, (1982, P. 100), with further clarification given in http://www.rasch.org/rmt/rmt142m.htm.

Usage
```r
pairwise.person.fit(pers_obj, na_treat = NA)
```

Arguments

- **pers_obj**: an object of class "pers" as a result from function `pers`
- **na_treat**: value to be assigned to residual cells which have missing data in the original response matrix. default is set to `na_treat=NA` to ignore these cells in further calculations. An option is to set these residuals to 0 using `na_treat=0`, which implies that they are imputed as 'fitting data', i.e., zero residuals. This can attenuate contrasts (see. http://www.rasch.org/rmt/rmt142m.htm).
pairwise.S

Details

contrary to many IRT software using ML based item parameter estimation, pairwise will not exclude persons, showing perfect response vectors (e.g. c(0,0,0) for dataset with three variables), prior to scaling. Therefore the fit statistics computed with pairwise may deviate somewhat from the fit statistics produced by IRT software using ML based item parameter estimation (e.g. R-package eRm), depending on the amount of persons with perfect response vectors in the data.

Value

an object of class c("pairwise_person_fit", "data.frame") containing person fit indices

Examples

############################################################
data(sim200x3)
result <- pers(pair(sim200x3))
pairwise.person.fit(pers.obj=result) # item fit statistic

Description

This function calculates the S-statistic on item level proposed by Fischer and Scheiblechner (1970) on item level for dicho- or polytomous item response formats by splitting the data into two subsamples. For polytomous Items the test is performed on item category level. Several splitting options are available (see arguments). The S-statistic is also mentioned in van den Wollenberg, (1982) – an article in Psychometrika, which might be available more easily (see details).

Usage

pairwise.S(daten, m = NULL, split = "random", splitseed = "no",
verbose = FALSE, ...)

Arguments

daten a data.frame or matrix with optionally named columns (names of items), potentially with missing values, comprising polytomous or dichotomous (or mixed category numbers) responses of n respondents (rows) on k items (columns) coded starting with 0 for lowest category to m-1 for highest category, with m being a vector (with length k) with the number of categories for the respective item.

m an integer (will be recycled to a vector of length k) or a vector giving the number of response categories for all items - by default m = NULL, m is calculated from data, assuming that every response category is at least once present in data. For sparse data it is strongly recommended to explicitly define the number of categories by defining this argument.
pairwise.S

split

Specifies the splitting criterion. Basically there are three different options available - each with several modes - which are controlled by passing the corresponding character expression to the argument.

1) Using the rawscore for splitting into subsamples with the following modes:
   split = "median" median raw score split - high score group and low score group; split = "mean" mean raw score split - high score group and low score group.
   2) Dividing the persons in daten into subsamples with equal size by random allocation with the following modes: split = "random" (which is equivalent to split = "random.2") divides persons into two subsamples with equal size. In general the number of desired subsamples must be expressed after the dot in the character expression - e.g. split = "random.6" divides persons into 6 subsamples (with equal size) by random allocation etc.
   3) The third option is using a manifest variable as a splitting criterion. In this case a vector with the same length as number of cases in daten must be passed to the argument grouping the data into subsamples. This vector should be coded as "factor" or a "numeric" integer vector with min = 1.

splitseed

numeric, used for set.seed(splitseed) for random splitting - see argument split.

verbose

logical, if verbose = TRUE (default) a message about subsampling is sent to console when calculating standard errors.

... additional arguments nsample, size, seed, pot for calling pairwise are passed through - see description for pairwise.

Details

The data is splitted in two subsamples and then item thresholds, the parameter (Sigma) and their standard errors (SE) for the items according the PCM (or RM in case of dichotominies) are calculated for each subsample. This function internally calls the function pairwise. Additional arguments (see description of function pairwise) for parameter calculation are passed through. This item fit statistic is also (perhaps misleadingly) named as 'Wald test' in other R-packages. The S-statistic, as implemented in pairwise, is defined according to Fischer and Scheiblechner (1970); see also equation (3) in van den Wollenberg, (1982), p. 124 in the following equation:

\[
S_i = \frac{\hat{\sigma}_i^{(1)} - \hat{\sigma}_i^{(2)}}{\sqrt{(S_{\hat{\sigma}_i}^{(1)})^2 + (S_{\hat{\sigma}_i}^{(2)})^2}}
\]

where \(\hat{\sigma}_i^{(1)}\) is the estimate of the item parameter of subsample 1, \(\hat{\sigma}_i^{(2)}\) is the estimate of the item parameter of subsample 2 and \(S_{\hat{\sigma}_i}^{(1)}\) and \(S_{\hat{\sigma}_i}^{(2)}\) are the respective standard errors. In Fischer (1974), p. 297, the resulting test statistic (as defined above) is labeled with \(Z_i\), as it is asymptotically normally distributed. Contrary to the 'Wald-type' test statistic \(W_i\), which was derived by Glas and Verhelst (2005) from the (general) \(\chi^2\) distributed test of statistical hypotheses concerning several parameters, which was introduced by Wald (1943).

Value

A (list) object of class "pair.S" containing the test statistic and item difficulty parameter sigma and their standard errors for the two or more subsamples.
A note on standard errors

Estimation of standard errors is done by repeated calculation of item parameters for subsamples of the given data. This procedure is mainly controlled by the arguments `nsample` and `size` (see arguments in `pairSE`). With regard to calculation time, the argument `nsample` is the 'time killer'. On the other hand, things (estimation of standard errors) will not necessarily get better when choosing large values for `nsample`. For example choosing `nsample=400` will only result in minimal change for standard error estimation in comparison to (`nsample=30`) which is the default setting (see examples).

References

description of function `pairSE(pairwise)`.


Examples

```r
### polytomous examples
data(bfiN) # loading example data set
data(bfi_cov) # loading covariates to bfiN data set

# calculating itemparameters and SE for two subsamples by gender
S_gen <- pairwise.S(daten=bfiN, split = bfi_cov$gender)
summary(S_gen)
summary(S_gen, thres = FALSE)

# other splitting criteria
## Not run:
S_med <- pairwise.S(daten=bfiN, split = "median")
summary(S_med)

S_ran<-pairwise.S(daten=bfiN, split = "random")
summary(S_ran)

S_ran.4<-pairwise.S(daten=bfiN, split = "random.4")
summary(S_ran.4) # currently not displayed
```
pers

### Example from details section 'Some Notes on Standard Errors'

```r
S_def <- pairwise.S(daten = bfiN, split = "random", splitseed = 13)
summary(S_def)
```

```r
S_400 <- pairwise.S(daten = bfiN, split = "random", splitseed = 13, nsample = 400)
summary(S_400)
```

## End (Not run)

---

**WLE - Rasch Person Parameter**

### Description

This is the (new) main function for calculation of person estimates based on answering dichotomous or polychotomous items according to Rasch Model (Rasch, 1960) and Partial Credit Model (Masters, 1982), given the item parameters (object of class "pair" - as a result of `pair()` and and the datamatrix (argument `daten`) containing the person response vectors (rows), using an WL approach, introduced by Warm (1989).

### Usage

```r
pers(itempar, daten = NULL, incidenz = NULL, na_treat = NULL,
     limit = 1e-05, iter = 50, Nrel = FALSE, tecout = FALSE)
```

### Arguments

- **itempar**: The item parameter prior calculated or estimated. A list object of class "pair" as a result of applying the function `pair()` to the data. Or an 'ordinary' "matrix" with `nrow = k` (number of items) and `ncol = m` (maximum number of thresholds), holding the 'thronsonian' thresholds of the respective item. Some matrix entries may be `NA`, depending on the number of categories of the respective item.

- **daten**: A "matrix" (or "data.frame") optionaly with named columns (names of items) and named rows (person IDs). This argument can be left empty when the argument `itempar` (above) is of class "pair". `daten` holds polychotomous or dichotomous (or mixed category numbers) responses of `n` respondents (rows) on `k` items (columns) coded starting with 0 for lowest category to `m-1` for highest category, with `m` being a vector (with length `k`) with the number of categories for the respective item. Responses in `daten` must be stored as "integers" (not "factors"!) and may have missing values.

- **incidenz**: This argument is only relevant when items are assigned to different booklets. For such a booklet-design a "matrix" should be assigned to this argument, with the same dimensions like `daten`, containing 0 and 1 integer codes, giving the information (for every person) if the respective item was in the respective booklet (coded 1) given to the person or not (coded 0).
na_treat  

Optionally an integer (vector) defining the type of treatment to missing responses in the argument `daten`. If set to `na_treat=NULL` (default) missing responses are treated as missings and the respective person is assigned to an corresponding missing group for estimation. An option is to set `na_treat` to any integer value between 0 (lowest category) and the numeric code for the maximum category of the respective item.

limit  

Numeric giving the limit at which accuracy the WL-algorithm stops.

iter  

Numeric giving the maximum number of iteration to perform.

nrel  

Logical with default set to `nrel=FALSE` to include persons with perfect response vectors for calculating WLE reliability. If set to `nrel=TRUE` persons with perfect response vectors are excluded for calculating WLE reliability.

tecout  

Logical default set to `FALSE`. If set to `TRUE` the result will be a (very) long list with estimation details for every case in `daten`. In case of a booklet-design the list entries will be divided by "booklet".

Details  

No detail in the moment.

Value  

An object of class `c("pers", "data.frame")` or a (very long) "list" (when setting on `techout=TRUE`) containing the person parameters.

References  


Examples  

```R
# Example
data(sim200x3)
result <- pers(itempar=pair(sim200x3))
summary(result)
plot(result)
logLik(result)  # Log-Likelihood for 'estimated' model
logLik(result, sat=TRUE)  # Log-Likelihood for saturated model
AIC(logLik(result))  # AIC for 'estimated' model
AIC(logLik(result, sat=TRUE))  # AIC for saturated model
BIC(logLik(result))  # BIC for 'estimated' model
BIC(logLik(result, sat=TRUE))  # BIC for saturated model

# Following example requires package eRm
# require(eRm)
# itemparameter with eRm:
# itempar_eRm <- thresholds(PCM(sim200x3))$ threshtable[[1]][,2:3]
```
## Description

S3 plotting Method for object of class "grm"

## Usage

```r
## S3 method for class 'grm'
plot(x, xlab = NULL, ylab = NULL, main = NULL,
     col.error = "blue", col.diag = "red", itemNames = TRUE,
     cex.names = 0.8, type = "b", xlab = NULL, ylab = NULL, pch = 43,
     las = 3, cex.axis = 0.5, ...)  
```

## Arguments

- **x**: object of class "grm"
- **xlab**: optional lower limit for xy-axis
- **ylab**: optional upper limit for xy-axis
- **ci**: numeric defining confidence intervall for point estimator
- **main**: see `plot`
- **col.error**: vector of colors for error bars
- **col.diag**: color for the diagonal of the plot
- **itemNames**: logical wether to plot itemnames
- **cex.names**: magnification factor for itemnames
- **type**: see `plot`
- **xlab**: see `plot`
- **ylab**: see `plot`
- **pch**: see `plot`
- **las**: see `plot`
- **cex.axis**: see `plot`
- **...**: other parameters passed to `plot`
plot.pair

S3 Plotting Thurstonian Thresholds

Description

S3 plotting Method for object of class "pair"

Usage

## S3 method for class 'pair'
plot(x, sortdif = FALSE, ra = "auto", main = NULL,
     col.lines = (1:dim(x$threshold)[2]), type = "b", xlab = "items",
     ylab = "logits", pch = (1:dim(x$threshold)[2]), las = 3,
     cex.axis = 0.8, ...) 

Arguments

x          object of class "pair"

sortdif    logical whether to order items by difficulty

ra         either the character "auto" (default) or an numeric, defining the (logit) range for y-axis

main       see plot

col.lines  vector of colors for threshold profile lines

type       see plot

xlab       see plot

ylab       see plot

pch        see plot

las         see plot

cex.axis   see plot

...        other parameters passed to plot

plot.pairSE

S3 Plotting Thurstonian Thresholds with SE

Description

S3 plotting method for object of class "pairSE"
Usage

```r
## S3 method for class 'pairSE'
plot(x, ci = 2, sortdif = FALSE, ra = "auto",
    main = NULL, col.lines = 1:(dim(x$threshold)[2]),
    col.error = 1:(dim(x$threshold)[2]), type = "b", xlab = "items",
    ylab = "logits", pch = 20, las = 3, cex.axis = 0.8, ...)
```

Arguments

- `x` object of class "pairSE"
- `ci` numeric defining confidence intervall for point estimator
- `sortdif` logical wether to order items by difficulty
- `ra` either the character "auto" (default) or an numeric, defining the (logit) range for y-axis
- `main` see `plot`
- `col.lines` vector of colors for threshold profile lines
- `col.error` vector of colors for error bars
- `type` see `plot`
- `xlab` see `plot`
- `ylab` see `plot`
- `pch` see `plot`
- `las` see `plot`
- `cex.axis` see `plot`
- `...` other parameters passed to `plot`

---

`plot.pers`  
**S3 Plotting Person - Item Map**

Description

S3 plotting method for object of class "pers"

Usage

```r
## S3 method for class 'pers'
plot(x, ra = NULL, sortdif = FALSE, main = NULL,
    ylab = "Logits", itemNames = TRUE, fillCol = "grey60",
    lineCol = "grey40", cex = 0.7, pos = 4, breaks = "Sturges", pch = 1,
    ...
```

Arguments

- **x**: object of class "pers"
- **ra**: an integer, defining the (logit) range for y-axis
- **sortdif**: logical whether to order items by difficulty
- **main**: see `plot`
- **ylab**: see `plot`
- **itemNames**: logical whether to use itemnames in the resulting plot
- **fillCol**: color for bar filling of the ability histogram
- **lineCol**: color for bar lines of the ability histogram
- **cex**: see `text`
- **pos**: see `text`
- **breaks**: see `hist`
- **pch**: see `points`
- **...**: other parameters passed to `hist` and `text`.

**plot.rfa**

*S3 Plotting Rasch Residual Factor Analysis*

Description

S3 plotting Method for object of class "rfa"

Usage

```r
## S3 method for class 'rfa'
plot(x, com = 1, ra = "auto", main = NULL, labels = NULL,
     xlab = "logits", ylab = "loadings", srt = 0, cex.axis = 0.8,
     cex.text = 0.8, col.text = NULL, ...)
```

Arguments

- **x**: object of class "rfa"
- **com**: an integer giving the number of the principal component used for plotting
- **ra**: either the character "auto" (default) or an numeric, defining the (logit) range for x-axis
- **main**: see `plot`
- **labels**: a character vector specifying the plotting pattern to use. see `text`. At default the itemnames are used.
- **xlab**: see `plot`
- **ylab**: see `plot`
- **srt**: see `text` or `par`
cex.axis see plot
cex.text see argument cex in function text
col.text see argument col in function text
... other parameters passed through.

---

**ptbis**

*Point Biserial Correlations*

**Description**
Calculation of the point biserial correlations for dicho- or polytomous item categories with total scale (person parameter).

**Usage**

```r
ptbis(y, daten = NULL)
```

**Arguments**

- `y` either an object of class "pers", or a numeric vector as a result of any scaling approach (WLE, MLE, Rawscore, etc.) relating to the Items (columns) in `daten`.
- `daten` if argument `y` is not an object of class "pers", a "data.frame", potentially with missing values, comprising dicho- or polytomous items (columns).

**Details**
no details in the moment.

**Value**
An object of class c("data.frame", "ptbis") containing item statistics.

**Examples**

```r
############################
#####
data(sim200x3) # loading response data
y <- rowSums(sim200x3)
ptbis(y=y, daten=sim200x3)
####
result <- pers(pair(sim200x3))
ptbis(y=result)
```
function for calculating the person fit index Q, which was proposed by Tarnai and Rost (1990).

Usage

Q(obj, ...)

Arguments

obj: an object of class "pers" or class "pair" as a result from function `pers` or `pair` respectively.

Details

The person Q-index proposed by Tarnai and Rost, (1990) is solely based on the empirical responses and the item parameters. Thus the computation of person parameters using the function `pers` is not required - see examples. But for convenience return objects of both functions are accepted in function Q.

Value

a vector holding the Q-index for every person.

References


Examples

```r
####
data(bf1N) # get some data
ip <- pair(daten = bf1N,m = 6) # item parameters according the partial credit model
Q(ip)
```
**q3**

**Q3 Fit Statistic**

**Description**
Calculation of Q3 fit statistic for the rasch model based on the residuals, which was proposed by Yen (1984).

**Usage**

```
q3(pers_obj, na_treat = 0, use = "complete.obs", res = "stdr",
   method = "pearson")
```

**Arguments**

- **pers_obj**: an object of class "pers" as a result from function `pers`.
- **na_treat**: value to be assigned to residual cells which have missing data in the original response matrix. default is set to `na_treat=0` to set the residuals to 0, which implies that they are imputed as 'fitting data', i.e., zero residuals. This can attenuate contrasts (see. http://www.rasch.org/rmt/rmt142m.htm). An option is to set it to `na_treat=NA`.
- **use**: a character string as used in function `cor` or `cov`, giving a method for computing covariances or correlations in the presence of missing values. This must be (an abbreviation of) one of the strings "everything", "all.obs", "complete.obs", "na.or.complete", or "pairwise.complete.obs". The default is set to `use="complete.obs"` which will exclude cases by listwise deletion to keep the correlation matrix positive definite.
- **res**: a character string defining which type of (rasch–) residual to analyze when computing the correlations. This must be (exactly) one of the strings "sr" for score residuals , "stdr" for standardised residuals, "srsq" for score residuals squared, or "stdrsq" for standardised residuals squared. The default is set to `res="stdr"` referring to Linacre (1998).
- **method**: a character string as used in function `cor`, indicating which correlation coefficient is to be computed. One of "pearson" (default), "kendall", or "spearman", can be abbreviated. The default is set to `method="pearson"`.

**Details**
The lower level letter 'q' was used (instead of 'Q') for naming the function because the name 'Q3' was already used in another IRT package – namely TAM. As perhaps some users like to use both packages simultaneously, an alternative naming convention was chosen for 'pairwise'. No other details in the moment.

**Value**

An object of class `c("Q3","list")`. 
References


Examples

```r
############
data(bfiN) # loading response data
pers_obj <- pers(pair(bfiN))
result <- q3(pers_obj)
str(result) # to see what's in ;-)  
```

---

**rfa**  
*Rasch Residual Factor Analysis*

**Description**

Calculation of the rasch residual factor analysis proposed by Wright (1996) and further discussed by Linacre (1998) to detect multidimensionality.

**Usage**

```r
rfa(pers_obj, na_treat = 0, tr = FALSE, use = "complete.obs", res = "stdr", method = "pearson", cor = TRUE)
```

**Arguments**

- `pers_obj` an object of class "pers" as a result from function `pers`.
- `na_treat` value to be assigned to residual cells which have missing data in the original response matrix. default is set to `na_treat=0` to set the residuals to 0, which implies that they are imputed as 'fitting data', i.e., zero residuals. This can attenuate contrasts (see. http://www.rasch.org/rmt/rmt142m.htm). An option is to set it to `na_treat=NA`.
- `tr` a logical value indicating whether the data (the residual matrix) is transposed prior to calculation. This would perform a person analysis rather than an item analysis. The default is set to item analysis.
- `use` a character string as used in function `cor` or `cov`, giving a method for computing covariances or correlations in the presence of missing values. This must be (an abbreviation of) one of the strings "everything", "all.obs", "complete.obs", "na.or.complete", or "pairwise.complete.obs". The default is set to `use="complete.obs"` which will exclude cases by listwise deletion to keep the correlation matrix positive definite.
res  a character string defining which type of (rasch-) residual to analyze when computing covariances or correlations. This must be (exactly) one of the strings "sr" for score residuals, "stdr" for standardised residuals, "srsq" for score residuals squared, or "stdrsq" for standardised residuals squared. The default is set to res="stdr" refering to Linacre (1998).

method  a character string as used in function cor or cov, indicating which correlation coefficient (or covariance) is to be computed. One of "pearson" (default), "kendall", or "spearman", can be abbreviated. The default is set to method="pearson".

cor  a logical value indicating whether the calculation should use the correlation matrix or the covariance matrix. The default is set to cor=TRUE to use the correlation matrix.

Details  no details in the moment.

Value  An object of class c("rfa","list").


Examples  

```
###########################
# data(bfiN)  # loading reponse data
pers_obj <- pers(pair(bfiN))
result <- rfa(pers_obj)
summary(result)
plot(result)
```

Simulated Data

Simulated data for 200 'subjects' 'answering' to 3 items with unequal number of categories – one dichotomous and two polytomous items.
**simra**

**Usage**

```r
data(sim200x3)
```

**Format**

A data.frame containing 3 variables and 200 observations.

**Details**

This simulated data is used as an example in the rasch module of the ’ALMO - Statistiksystem’.

**Source**

http://www.almo-statistik.de/

**References**


**Examples**

```r
data(sim200x3)
dim(sim200x3)
apply(sim200x3,2,table)
```

---

**simra**

*Simulate Response Pattern under Dichotomous and Polytomous Rasch Model*

**Description**

function for simulation of response patterns following the dichotomous and/or polynomous Rasch model based on the category probabilities given the model parameters. At default, when just calling `simra()` 1 replication of responses to 5 items with difficulties -2, -1, 0, 1, 2 from 100 persons with ability drawn from N(0|1) are sampled.

**Usage**

```r
simra(itempar = matrix(seq(-2, 2, length = 5)), theta = 100,
    pers_obj = NULL, replicate = 1, seed = seq(1, replicate, 1), ...)
```
Arguments

itempar a "matrix" with nrow = k (number of items) and ncol = m (maximum number of thresholds), holding the 'thurstonian' thresholds of the respective item. Some of the rightmost matrix entries may be NA, depending on the number of categories of the respective item.

theta either one of the following (1) a numeric vector of length n providing the values of the person parameter (ability) for n persons to be used in the simulation. (2) a integer defining the number of values n to draw from N(0|1).

pers_obj an object of class "pers" as a result from function pers. If an object of class "pers" is assigned to this argument the model parameters in it are taken to simulate the responses. At default (pers_obj = NULL) simulation is done by considering the model parameters given in the other arguments (see above).

replicate an integer defining how many replicates (data matrices) r to draw based on the model parameters.

seed a numeric vector with legnth of number of replications used for set.seed prior to each replicate to keep the result repeatable. If seed = NULL no seed is set.

... arguments passed through.

Details

no details in the moment.

Value

an array with dim(n,k,r) response patterns (k items in columns n persons in rows and r replications in the third dimension).

Examples

# 100 dichotomous probabilistic response pattern
# 100 polytomous response pattern (4 items; each 4 answer categories)
v <- c(-1.0,-0.5,0.0,0.5,-0.75,-0.25,0.25,0.75,-0.5,0.0,0.5,1.0)
itempar <- matrix(v,nrow = 4,ncol = 3)
simra(itempar)
simra(itempar = itempar,replicate = 10) # draw 10 replications

summary.grm S3 Summary for graphical Model Check

Description

S3 summary method for object of class "grm"
Summary

**Usage**

```r
## S3 method for class 'grm'
summary(object, ci = 2, ...)
```

**Arguments**

- `object`: object of class "grm"
- `ci`: numeric with default `ci=2` to return confidence intervalls for point estimator.
- `...`: other parameters passed trough

**Description**

S3 summary method for object of class "grm"

**Usage**

```r
## S3 method for class 'pair'
summary(object, sortdif = FALSE, ...)
```

**Arguments**

- `object`: object of class "pair"
- `sortdif`: logical with default `sortdif=FALSE` wether to order items by difficulty.
- `...`: other parameters passed trough

**Summary.pair.S**

**S3 Summary for S-statistic Test (Wald Test)**

**Description**

S3 summary method for object of class "pair.S"

**Usage**

```r
## S3 method for class 'pair.S'
summary(object, thres = TRUE, ...)
```

**Arguments**

- `object`: object of class "pair.S"
- `thres`: logical whether to output results based on the thresholds
- `...`: other parameters passed trough
**summary.pairSE**

*S3 Summary for item parameter with standard errors*

**Description**

S3 summary method for object of class "pairSE"

**Usage**

```r
## S3 method for class 'pairSE'
summary(object, sortdif = FALSE, ...)
```

**Arguments**

- `object` : object of class "pairSE"
- `sortdif` : logical with default sortdif=FALSE wether to order items by difficulty.
- `...` : other parameters passed trough

**summary.pairwise.item.fit**

*S3 Summary for Item-Fit-Statistics*

**Description**

S3 summary method for object of classc("pairwise.item.fit", "data.frame" )

**Usage**

```r
## S3 method for class 'pairwise.item.fit'
summary(object, sort = FALSE, by = "INFIT.ZSTD",
         decreasing = FALSE, ...)
```

**Arguments**

- `object` : object of class "pairwise.item.fit", "data.frame"
- `sort` : logical with default sort=FALSE - if set to sort=TRUE items are ordered by absolute FIT.
- `by` : character passing the type of Fit-Statistic to sort by - ignored when sort=FALSE. valid options are: "INFIT.ZSTD" (default), "OUTFIT.MSQ", "OUTFIT.ZSTD" and "INFIT.MSQ".
- `decreasing` : see order
- `...` : other parameters passed trough - see order
summary.pairwise.person.fit

S3 Summary for Person-Fit-Statistics

Description

S3 summary method for object of class("pairwise.item.fit", "data.frame")

Usage

## S3 method for class 'pairwise.person.fit'
summary(object, sort = FALSE,
       by = "INFIT.ZSTD", decreasing = FALSE, ...)

Arguments

object object of class"pairwise.person.fit", "data.frame"
sort logical with default sort=FALSE - if set to sort=TRUE persons are ordered by
       absolute FIT.
by character passing the type of Fit-Statistic to sort by - ignored when sort=FALSE. valid options are: "INFIT.ZSTD" (default), "OUTFIT.MSQ","OUTFIT.ZSTD" and
"INFIT.MSQ".
decreasing see order
... other parameters passed trough - see order

summary.pers

S3 Summary for Thetas

Description

S3 summary method for object of class"pers"

Usage

## S3 method for class 'pers'
summary(object, short = TRUE, sortwle = FALSE, ...)

Arguments

object object of class"pers"
short logical with default short=TRUE - if set to short=FALSE a "data.frame" with
       WLE estimates (and their respective standard errors) for every row (person) in
       the original dataset will be returned.
sortwle logical wether to order persons by ability - ignored when short=TRUE
... other parameters passed trough
**summary.rfa**  
*S3 Summary for rasch factor analysis*

**Description**
S3 summary method for object of class "pair"

**Usage**
## S3 method for class 'rfa'
summary(object, sortdif = FALSE, ...)

**Arguments**
- **object**: object of class "pair"
- **sortdif**: logical with default sortdif=FALSE wether to order items / persons by difficulty / ability.
- **...**: other parameters passed trough

**tff**  
*Test information function*

**Description**
plotting function for plotting the test information function (TIF).

**Usage**
tff(pair_obj, items = NULL, x = NULL, main = "Test Information Function", plot = TRUE, cat = FALSE, lwd = 2, col = 1, ...)

**Arguments**
- **pair_obj**: an object of class "pair" as a result from function pair.
- **items**: optional a vector (character or numeric) identifying the items (according their order in the data) to use for plotting the test information function.
- **x**: The value(s) of the latent variable, at which the TIF will be evaluated. x should be either a numeric vector of theta values or a single numeric value. If x is given as a single numeric value plotting is supressed. If not given (default), 99 values spaced evenly between -4 and +4 will be used, handy for plotting.
- **main**: see parameters for plot
- **plot**: a logical (default plot = TRUE), defining wether to supress plotting an just return a matrix of the values of the Item information function.
cat  a logical (default cat = FALSE), defining whether to plot as an overlay to the Test information function the item category information functions based on item categories. If cat = TRUE and plot = FALSE the values of the item category information functions are returned.

lwd  see parameters for plot

col  see parameters for plot

... arguments passed to plot

Details
	ono details in the moment.

Value

a plot, a "data.frame" or a single numeric with values of the Test information function.

Examples

```
  #########
  data(sim200x3)
  result <- pair(sim200x3)
  tff(pair_obj = result) # TIF plot
  tff(pair_obj = result, cat=TRUE) # TIF plot
  tff(pair_obj = result, items=c("V1","V3"), cat=TRUE) # TIF plot
  tff(pair_obj = result, x=0) # TIF at theta=0
  tff(pair_obj = result, x=seq(0,4,.1)) # TIF for a given range of Thetas
  ######## examples with other data ...
  data(bfiN)
  result <- pair(bfiN)
  tff(pair_obj = result)
  tff(pair_obj = result, cat=TRUE) # TIF plot
```
Index

+Topic datasets
  bfi_cov, 9
  bfin, 6
  bfin_miss, 8
cog, 11
cogBOOKLET, 12
DEU_PISA2012, 14
kft5, 21
Neoffi5, 24
sim200x3, 44

andersentest.pers, 5

bfi_cov, 9
bfin, 6, 9
bfin_miss, 8, 9
catprob, 10, 26
cog, 11, 12
cogBOOKLET, 12, 23
cor, 42–44
cov, 42–44
deltapar, 12
DEU_PISA2012, 14
esc, 4, 14
ftab, 15
gif, 3, 16
grm, 4, 5, 17
hist, 39
iff, 4, 20
kft5, 21
logLik.pers, 22
lrtest.pers, 23

make.incidenz, 3, 23
Neoffi5, 24
order, 48, 49
pair, 3, 5, 10, 12, 13, 20, 25, 34, 41, 50
pairSE, 17–19, 27, 32, 33
pairwise (pairwise-package), 3
pairwise-package, 3
pairwise.item.fit, 3, 29
pairwise.person.fit, 3, 30
pairwise.S, 31
par, 39
ers, 3, 5, 15, 16, 22, 23, 29, 30, 34, 41–43, 46
plot, 15, 20, 36–40, 50, 51
plot.grm, 4, 36
plot.pair, 26, 37
plot.pairSE, 37
plot.pers, 38
plot.rfa, 39
points, 39
ptbis, 3, 40

Q, 41
q3, 42
rfa, 4, 43

set.seed, 46
sim200x3, 44
simra, 45
summary.grm, 46
summary.pair, 47
summary.pair.S, 47
summary.pairSE, 48
summary.pairwise.item.fit, 48
summary.pairwise.person.fit, 49
summary.pers, 49
summary.rfa, 50
INDEX

text, 39, 40
tff, 4, 50