

Package ‘PROsetta’

July 1, 2020

Type Package

Title Linking Patient-Reported Outcomes Measures

Version 0.1.4

Date 2020-06-30

Description Perform scale linking to establish relationships between instruments that measure similar constructs according to the PROsetta Stone methodology, as in Choi, Schalet, Cook, & Cella (2014) <doi:10.1037/a0035768>.

URL <http://prosettastone.org>

BugReports <https://github.com/choi-phd/PROsetta/issues>

Imports equate, lavaan, mirt, plink, psych, methods

Depends R (>= 3.5.0)

Suggests shiny, shinythemes, shinyWidgets, shinyjs, DT, knitr, kableExtra, testthat (>= 2.1.0), dplyr, rmarkdown, pkgdown

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

VignetteBuilder knitr

Collate 'import.R' 'configPROsetta.R' 'post_functions.R'
'core_functions.R' 'datasets.R' 'example.R'
'linking_functions.R' 'plot_functions.R' 'runshiny.R'

NeedsCompilation no

Author Seung W. Choi [aut, cre] (<<https://orcid.org/0000-0003-4777-5420>>),
Sangdon Lim [aut] (<<https://orcid.org/0000-0002-2988-014X>>),
Benjamin D. Schalet [ctb],
Aaron Kaat [ctb],
David Cella [ctb]

Maintainer Seung W. Choi <schoi@austin.utexas.edu>

Repository CRAN

Date/Publication 2020-07-01 11:00:18 UTC

R topics documented:

checkFrequency	2
compareScores	3
dataset_asq	3
dataset_dep	4
getCompleteData	5
getEscore	6
getScaleSum	6
getTheta	7
loadData	8
plot,PROsetta_data,ANY-method	9
plotInfo	10
PROsetta	11
runCalibration	11
runCFA	12
runClassical	13
runDescriptive	14
runEquateObserved	14
runFrequency	16
runLinking	17
runRSSS	18

Index

19

checkFrequency

Check frequency table for unobserved response categories

Description

checkFrequency is a descriptive function to check whether all response categories in a frequency table have a frequency of at least 1.

Usage

```
checkFrequency(data)
```

Arguments

data a **PROsetta_data** object. See **loadData** for loading a dataset.

Value

If all response categories have a frequency of at least 1, the value is TRUE.

Otherwise, the value is FALSE.

compareScores	<i>Compare two sets of scores</i>
---------------	-----------------------------------

Description

`compareScores` is a helper function to compare two sets of scores.

Usage

```
compareScores(left, right, type = c("corr", "mean", "sd", "rmsd", "mad"))
```

Arguments

- | | |
|-------|---|
| left | scores on the left side of comparison. |
| right | scores on the right side of comparison. This is subtracted from 'left'. |
| type | type of comparisons to include. Accepts 'corr', 'mean', 'sd', 'rmsd'. Defaults to all four. |

Value

`compareScores` returns a `data.frame` containing the comparison results.

dataset_asq	<i>ASQ dataset</i>
-------------	--------------------

Description

This dataset is associated with the following objects:

Details

- `response_asq` a `data.frame` containing raw response data of 751 participants and 41 variables.
 - prosettaid. participant IDs.
 - EDANX01 --MASQ11. response to items.
- `itemmap_asq` a `data.frame` containing the item map, describing the items in each instrument.
 - `item_order` item numeric IDs. This column refers to the column `item_order` in anchor item attributes.
 - `instrument` the instrument ID that each item belongs to.
 - `item_id` item ID strings. This column refers to column names in raw response data, excluding the participant ID column.
 - `item_name` new item ID strings to be used in the combined scale.
 - `ncat` the number of response categories.

- `min_score` the minimum score of the item.
- `reverse` whether the item data has been reverse-scored. 1 indicates the item has been reverse-scored, and 0 indicates the item has not been reverse-scored.
- `scores` a comma-separated string representing all possible score values from the item.
- `anchor_asq` a `data.frame` containing anchor item parameters for 29 items.
 - `item_order` item numeric IDs.
 - `item_id` item ID strings. This column refers to column names in raw response data, excluding the participant ID column.
 - `a` the discrimination (slope) parameter for the graded response model.
 - `cb1 -cb4` the boundaries between each category-pair for the graded response model.
 - `ncat` the number of response categories.
- `data_asq` a `PROsetta_data` object containing the datasets above. See `loadData` for creating `PROsetta_data` objects.

Examples

```
## load datasets into a PROsetta_data object
data_asq <- loadData(
  response = response_asq,
  itemmap = itemmap_asq,
  anchor = anchor_asq
)

## run descriptive statistics
runDescriptive(data_asq)

## run item parameter calibration on the response data, linking to the anchor item parameters
runLinking(data_asq, method = "FIXEDPAR")
```

dataset_dep

DEP dataset

Description

This dataset is associated with the following objects:

Details

- `response_dep` a `data.frame` containing raw response data of 747 participants and 49 variables.
 - `prosettaid`. participant IDs.
 - `EDDEP04 --CESD20`. response to items.
- `itemmap_dep` a `data.frame` containing the item map, describing the items in each instrument.
 - `item_order` item numeric IDs. This column refers to the column `item_order` in anchor item parameters.

- instrument the instrument ID that each item belongs to.
- item_id item ID strings. This column refers to column names in raw response data, excluding the participant ID column.
- item_name new item ID strings to be used in the combined scale.
- ncat the number of response categories.
- min_score the minimum score of the item.
- reverse whether the item data has been reverse-scored. 1 indicates the item has been reverse-scored, and 0 indicates the item has not been reverse-scored.
- scores a comma-separated string representing all possible score values from the item.
- anchor_dep a `data.frame` containing anchor item parameters for 28 items.
 - item_order item numeric IDs.
 - item_id item ID strings. This column refers to column names in raw response data, excluding the participant ID column.
 - a the discrimination (slope) parameter for the graded response model.
 - cb1 –cb4 the boundaries between each category-pair for the graded response model.
 - ncat the number of response categories.
- data_dep a `PROsetta_data` object containing the datasets above. See `loadData` for creating `PROsetta_data` objects.

Examples

```
## load datasets into a PROsetta_data object
data_dep <- loadData(
  response = response_dep,
  itemmap = itemmap_dep,
  anchor = anchor_dep
)

## run descriptive statistics
runDescriptive(data_dep)

## run item parameter calibration on the response data, linking to the anchor item parameters
runLinking(data_dep, method = "FIXEDPAR")
```

`getCompleteData` *Get complete data*

Description

`getCompleteData` is a helper function to perform casewise deletion of missing values.

Usage

```
getCompleteData(data, scale = NULL)
```

Arguments

- `data` a `PROsetta_data` object.
`scale` the index of the scale to perform casewise deletion. Leave empty or set to "combined" to perform on all scales.
-

<code>getEscore</code>	<i>Calculate expected scores at theta</i>
------------------------	---

Description

`getEscore` is a helper function to calculate expected scores at supplied thetas.

Usage

```
getEscore(ipar, model, theta, is_minscore_0)
```

Arguments

- `ipar` item parameters.
`model` item model to use.
`theta` theta values.
`is_minscore_0` if TRUE the score begins from 0 instead of 1.

Value

`getEscore` returns a vector of expected scores.

<code>getScaleSum</code>	<i>Calculate raw sum scores of a scale</i>
--------------------------	--

Description

`getScaleSum` is a helper function to calculate raw sum scores of a scale.

Usage

```
getScaleSum(data, scale_idx)
```

Arguments

- `data` a `PROsetta_data` object.
`scale_idx` the index of the scale to obtain the raw sum scores.

getTheta	<i>Obtain EAP estimates</i>
----------	-----------------------------

Description

`getTheta` is a helper function to calculate EAP estimates.

Usage

```
getTheta(  
  data,  
  ipar,  
  scale = "combined",  
  model = "grm",  
  theta_grid = seq(-4, 4, 0.1),  
  prior_dist = "normal",  
  prior_mean = 0,  
  prior_sd = 1  
)
```

Arguments

data	a <code>PROsetta_data</code> object.
ipar	a <code>data.frame</code> containing item parameters.
scale	the index of the scale to use. Set to 'combined' to use the combined scale.
model	the item model to use. Accepts 'grm' or 'gpcm'.
theta_grid	the theta grid to use in calculating EAP estimates.
prior_dist	the type of prior distribution. Accepts 'normal' or 'logistic'.
prior_mean	mean of the prior distribution.
prior_sd	SD of the prior distribution.

Value

`getTheta` returns a `list` containing EAP estimates.

<code>loadData</code>	<i>Load data from supplied config</i>
-----------------------	---------------------------------------

Description

`loadData` is a data loading function to create a `PROsetta_data` object, for scale linking/equating with 'PROsetta' package.

Usage

```
loadData(
  response,
  itemmap,
  anchor,
  item_id = NULL,
  person_id = NULL,
  scale_id = NULL,
  input_dir = getwd()
)
```

Arguments

<code>response</code>	response data containing case IDs and item responses. This can be a .csv filename or a <code>data.frame</code> object.
<code>itemmap</code>	an item map containing item IDs and scale IDs. This can be a .csv filename or a <code>data.frame</code> object.
<code>anchor</code>	anchor data containing item parameters for anchor items. This can be a .csv filename or a <code>data.frame</code> object.
<code>item_id</code>	the column name to look for item IDs. Automatically determined if not specified.
<code>person_id</code>	the column name to look for case IDs. Automatically determined if not specified.
<code>scale_id</code>	the column name to look for scale IDs. Automatically determined if not specified.
<code>input_dir</code>	the directory to look for the files.

Value

`loadData` returns a `PROsetta_data` object containing the loaded data.

```
plot,PROsetta_data,ANY-method
  Plot frequency distribution
```

Description

This is an extension of [plot](#) to visualize frequency distribution from [PROsetta_data](#) object.

Usage

```
## S4 method for signature 'PROsetta_data,ANY'
plot(
  x,
  y,
  scale_id = "combined",
  filename = NULL,
  title = NULL,
  xlim = NULL,
  color = "blue",
  nbar = 20,
  rug = FALSE,
  filetype = "pdf",
  savefile = FALSE,
  bg = "white",
  width = 6,
  height = 6,
  pointsize = 12
)
```

Arguments

x	a PROsetta_data object.
y	unused argument, exists for compatibility with plot in the base R package.
scale_id	scale ID to plot. <code>combined</code> (default) represents the combined scale.
filename	filename to write if 'savefile' argument is TRUE.
title	the title of the figure.
xlim	the range of scores to plot.
color	the color to fill the histogram.
nbar	the number of histogram bars.
rug	if TRUE, display the actual distribution of scores below each bar.
filetype	the type of file to write if 'savefile' argument is TRUE. Accepts 'pdf', 'jpeg', 'png', and 'tiff'.
savefile	if TRUE, save the figure as a file.
bg	the background color of the plot.

width the width of the plot.
height the height of the plot.
pointsize point size to pass onto file writing functions.

Examples

```
plot(data_asq)
plot(data_asq, scale_id = 1)
plot(data_asq, scale_id = 2)
```

plotInfo *Plot scale information*

Description

[plotInfo](#) is a plotting function to visualize scale-level information.

Usage

```
plotInfo(
  object,
  data,
  theta = seq(-4, 4, 0.1),
  t_score = FALSE,
  scale_label = c(1, 2, "Combined"),
  color = c("red", "blue", "black"),
  lty = c(3, 2, 1)
)

## S4 method for signature 'SingleGroupClass'
plotInfo(
  object,
  data,
  theta = seq(-4, 4, 0.1),
  t_score = FALSE,
  scale_label = c(1, 2, "Combined"),
  color = c("red", "blue", "black"),
  lty = c(3, 2, 1)
)
```

Arguments

object a [SingleGroupClass](#) object from [runCalibration](#).
data a [PROsetta_data](#) object.
theta theta values to plot on the x-axis.

t_score	set to TRUE to convert thetas into T-scores.
scale_label	names of each scale.
color	line colors to plot.
lty	line types to plot.

Examples

```
out_calib = runCalibration(data_asq, technical = list(NCYCLES = 1000))
plotInfo(out_calib, data_asq)
```

PROsetta

PROsetta

Description

PROsetta is a caller function to launch a Shiny app locally.

Usage

```
PROsetta()
guiPROsetta()
```

Examples

```
if (interactive()) {
  PROsetta()
}
```

runCalibration

Run Calibration

Description

runCalibration is a function to perform item calibration on the response data.

Usage

```
runCalibration(data, fixedpar = FALSE, ignore_nonconv = FALSE, ...)
```

Arguments

- data a PROsetta_data object. See [loadData](#) for loading a dataset.
- fixedpar if TRUE (default), perform fixed parameter calibration using anchor data.
- ignore_nonconv if TRUE, return results even when calibration did not converge. Defaults to FALSE.
- ... additional arguments to pass onto [mirt](#) in '[mirt](#)' package.

Value

[runCalibration](#) returns a [SingleGroupClass](#) object containing item calibration results.

This object can be used in [coef](#), [itemfit](#), [itemplot](#) in '[mirt](#)' package to extract wanted information.

Examples

```
## Not run:
out_calib <- runCalibration(data_asq) # errors

## End(Not run)

out_calib <- runCalibration(data_asq, technical = list(NCYCLES = 1000))

mirt::coef(out_calib, IRTpars = TRUE, simplify = TRUE)
mirt::itemfit(out_calib, empirical.plot = 1)
mirt::itemplot(out_calib, item = 1, type = "info")
mirt::itemfit(out_calib, "S_X2", na.rm = TRUE)
```

runCFA

Run a confirmatory factor analysis

Description

[runCFA](#) is a function to perform a one-factor confirmatory factor analysis (CFA) to test unidimensionality.

Usage

```
runCFA(data, estimator = "WLSMV", std.lv = TRUE, scalewise = FALSE, ...)
```

Arguments

- data a PROsetta_data object. See [loadData](#) for loading a dataset.
- estimator the estimator to be used. Passed onto [cfa](#) in '[lavaan](#)' package.
- std.lv if TRUE, the metric of the latent variable is determined by fixing their (residual) variances to 1.0. If FALSE, the metric of each latent variable is determined by fixing the factor loading of the first indicator to 1.0. Passed onto [cfa](#).

- scalewise if TRUE, run analysis for each scale as well as for the combined scale. If FALSE (default), run analysis only for the combined scale.
- ... additional arguments to pass onto `cfa`.

Value

`runCFA` returns a list containing the CFA results.

Examples

```
out_cfa <- runCFA(data_asq, scalewise = TRUE)
lavaan::summary(out_cfa$`1`      , fit.measures = TRUE, standardized = TRUE, estimates = FALSE)
lavaan::summary(out_cfa$`2`      , fit.measures = TRUE, standardized = TRUE, estimates = FALSE)
lavaan::summary(out_cfa$`combined`, fit.measures = TRUE, standardized = TRUE, estimates = FALSE)
```

runClassical

*Run CTT-based reliability analysis***Description**

`runClassical` is a function to perform Classical Test Theory (CTT) based reliability analysis.

Usage

```
runClassical(data, omega = FALSE, scalewise = TRUE, ...)
```

Arguments

- data a `PROsetta_data` object. See `loadData` for loading a dataset.
- omega if TRUE, also obtain McDonald's omega using `omega` in `psych` package. (default = FALSE)
- scalewise if TRUE, run analysis for each scale as well as for the combined scale. If FALSE, run analysis only for the combined scale. (default = TRUE)
- ... additional arguments to pass onto `omega`.

Value

`runClassical` returns a `list` containing reliability analysis results.

Examples

```
out_alpha <- runClassical(data_asq)
out_omega <- runClassical(data_asq, omega = TRUE) # also obtain omega
```

runDescriptive	<i>Obtain a descriptive statistics table</i>
----------------	--

Description

`runDescriptive` is a descriptive function to obtain descriptive statistics for each item in the dataset.

Usage

```
runDescriptive(data = NULL)
```

Arguments

data	a PROsetta_data object. See <code>loadData</code> for loading a dataset.
------	--

Value

`runDescriptive` returns a `data.frame` containing descriptive statistics (mean, standard deviation, median, ...) of the items in the dataset. These are calculated with `describe` in 'psych' package.

Examples

```
out_desc <- runDescriptive(data_asq)
```

runEquateObserved	<i>Run Test Equating</i>
-------------------	--------------------------

Description

`runEquateObserved` is a function to perform equipercentile test equating between two scales. A concordance table is produced, mapping the observed raw scores from one scale to the scores from another scale.

Usage

```
runEquateObserved(
  data,
  scale_from = 2,
  scale_to = 1,
  type_to = "raw",
  rsss = NULL,
  eq_type = "equipercentile",
  smooth = "loglinear",
  degrees = list(3, 1),
  boot = TRUE,
```

```
reps = 100,
...
)
```

Arguments

data	a PROsetta_data object. See loadData for loading a dataset.
scale_from	the scale ID of the input scale. References to itemmap in data argument.
scale_to	the scale ID of the target scale to equate to. References to itemmap in data argument.
type_to	the type of score to use in the target scale frequency table. Accepts raw, tscore, and theta. tscore and theta require argument rsss to be supplied. (default = raw)
rsss	the RSSS table to use to map each raw score level onto a t-score or a theta. See runRSSS .
eq_type	the type of equating to be passed onto equate in ' equate ' package.
smooth	the type of smoothing method to be passed onto presmoothing in ' equate ' package.
degrees	the degrees of smoothing to be passed onto presmoothing .
boot	performs bootstrapping if TRUE.
reps	the number of replications to perform in bootstrapping.
...	other arguments to pass onto equate .

Value

`runEquateObserved` returns an [equate](#) object containing the test equating result.

The printed summary statistics indicate the distributional properties of the two supplied scales and the equated scale.

- x corresponds to `scale_from`.
- y corresponds to `scale_to`.
- yx corresponds to `scale_from` after equating to `scale_to`.

See [equate](#) for details.

The concordance table is stored in `concordance` slot.

Examples

```
out_eq_raw <- runEquateObserved(data_asq,
  scale_to = 1, scale_from = 2,
  eq_type = "equipercentile", smooth = "loglinear"
)
out_eq_raw$concordance

out_link <- runLinking(data_asq, method = "FIXEDPAR")
```

```

out_rsss <- runRSSS(data_asq, out_link)
out_eq_tscore <- runEquateObserved(data_asq,
  scale_to = 1, scale_from = 2,
  type_to = "tscore", rsss = out_rsss,
  eq_type = "equipercentile", smooth = "loglinear"
)
out_eq_tscore$concordance

```

runFrequency*Obtain a frequency table***Description**

`runFrequency` is a descriptive function to obtain a frequency table from the dataset.

Usage

```
runFrequency(data, check_frequency = TRUE)
```

Arguments

- | | |
|------------------------------|---|
| <code>data</code> | a PROsetta_data object. See loadData for loading a dataset. |
| <code>check_frequency</code> | Logical. If TRUE, check the frequency table for missing response categories, and display warning message if any is missing. |

Value

`runFrequency` returns a [data.frame](#) containing the frequency table.

Examples

```

freq_asq <- runFrequency(data_asq)
freq_dep <- runFrequency(data_dep)

```

`runLinking`*Run Scale Linking*

Description

`runLinking` is a function to obtain item parameters from the response data, and perform scale linking onto the metric of supplied anchor item parameters.

Usage

```
runLinking(data, method, ...)
```

Arguments

`data` a `PROsetta_data` object. See `loadData` for loading a dataset.

`method` the type of linking to perform. Accepts:

- `MM` for mean-mean
- `MS` for mean-sigma
- `HB` for Haebara method
- `SL` for Stocking-Lord method
- `FIXEDPAR` for fixed parameter calibration

Linear transformation methods are performed with `plink` in '`plink`' package.

`...` additional arguments to pass onto `mirt` in '`mirt`' package.

Value

`runLinking` returns a `list` containing the scale linking results.

- `constants` linear transformation constants. NA if `method` argument was `FIXEDPAR`.
- `ipar_linked` item parameters calibrated to the response data, and linked to the anchor item parameters.
- `ipar_anchor` anchor item parameters used in linking.

Examples

```
out_link <- runLinking(data_asq, "SL", technical = list(NCYCLES = 1000))
out_link$constants # transformation constants
out_link$ipar_linked # item parameters linked to anchor
out_link <- runLinking(data_asq, "FIXEDPAR")
out_link$ipar_linked # item parameters linked to anchor
```

runRSSS*Run Crosswalk Table Generation***Description**

runRSSS is a function to generate raw-score to standard-score crosswalk tables from supplied calibrated item parameters.

Usage

```
runRSSS(
  data,
  ipar_linked,
  prior_mean = 0,
  prior_sd = 1,
  min_theta = -4,
  max_theta = 4,
  inc = 0.01,
  min_score = 1
)
```

Arguments

data	a PROsetta_data object. See loadData for loading a dataset.
ipar_linked	an object returned from runLinking or runCalibration .
prior_mean	prior mean.
prior_sd	prior standard deviation.
min_theta	the lower limit of theta grid.
max_theta	the upper limit of theta grid.
inc	the increment to use in theta grid.
min_score	minimum item score (0 or 1).

Value

runRSSS returns a [list](#) containing crosswalk tables.

Examples

```
out_link    <- runLinking(data_asq, method = "FIXEDPAR")
score_table <- runRSSS(data_asq, out_link)
```

Index

* datasets
 dataset_asq, 3
 dataset_dep, 4

anchor_asq, 4
anchor_asq (dataset_asq), 3
anchor_dep, 5
anchor_dep (dataset_dep), 4

cfa, 12, 13
checkFrequency, 2, 2
coef, 12
compareScores, 3, 3

data.frame, 3–5, 7, 8, 14, 16
data_asq, 4
data_asq (dataset_asq), 3
data_dep, 5
data_dep (dataset_dep), 4
dataset_asq, 3
dataset_dep, 4
describe, 14

equate, 15

getCompleteData, 5, 5
getEscore, 6, 6
getScaleSum, 6, 6
getTheta, 7, 7
guiPROsetta (PROsetta), 11

itemfit, 12
itemmap_asq, 3
itemmap_asq (dataset_asq), 3
itemmap_dep, 4
itemmap_dep (dataset_dep), 4
itemplot, 12

list, 7, 13, 17, 18
loadData, 2, 4, 5, 8, 8, 12–18

mirt, 12, 17

omega, 13

plink, 17
plot, 9
plot, PROsetta_data, ANY-method, 9
plotInfo, 10, 10
plotInfo, SingleGroupClass-method
 (plotInfo), 10
presmoothing, 15
PROsetta, 11, 11
PROsetta_data, 2, 4–10, 12–18
PROsetta_data-class (loadData), 8

response_asq, 3
response_asq (dataset_asq), 3
response_dep, 4
response_dep (dataset_dep), 4
runCalibration, 10, 11, 11, 12, 18
runCFA, 12, 12, 13
runClassical, 13, 13
runDescriptive, 14, 14
runEquateObserved, 14, 14, 15
runFrequency, 16, 16
runLinking, 17, 17, 18
runRSSS, 15, 18, 18

SingleGroupClass, 10, 12