

Package ‘dexter’

February 8, 2017

Type Package

Title Data Management and Basic Analysis of Tests

Version 0.1.7

Author Gunter Maris, Timo Bechger, Ivailo Partchev

Maintainer Ivailo Partchev <Ivailo.Partchev@cito.nl>

Description Data handling, Rasch model and Haberman's interaction model for educational and psychological tests that may involve multiple test forms or stages.

License GPL (>= 2)

Encoding UTF-8

LazyLoad yes

LazyData yes

Depends R (>= 2.10), RSQLite, plyr, reshape2, colorspace

Imports graphics, utils

RoxygenNote 5.0.1

Suggests knitr, rmarkdown, latticeExtra

VignetteBuilder knitr

NeedsCompilation yes

Repository CRAN

Date/Publication 2017-02-08 12:46:22

R topics documented:

add_booklet	2
add_item_properties	3
add_tests	4
distractor_plot	5
fit_domains	5
fit_models	6
keys_to_rules	7

open_project	8
PISA_item_class	8
plot.rim	9
print.rim	10
profile_plot	10
show_booklets	11
show_items	11
show_item_properties	12
show_person_properties	12
start_new_project	13
tia_tables	14
verbAggrData	14
verbAggrProperties	15
verbAggrRules	15

Index	16
--------------	-----------

add_booklet	<i>Add a booklet to a project</i>
-------------	-----------------------------------

Description

Adds item response data for a test form (a.k.a. booklet)

Usage

```
add_booklet(db, x, booklet_id)
```

Arguments

db	A handle to the database, i.e. the output of <code>start_new_project</code> or <code>open_project</code>
x	A data frame containing the responses and, possibly, some additional person characteristics. See details.
booklet_id	A (short) string identifying the test form (booklet)

Details

It is common practice to keep data in rectangular tables: data frames or foreign software like Excel, SPSS, etc. This function is provided to input data in that form, one booklet at a time. The starting point is a data frame, and getting the data frame into R is left to the user. We have found package `readxl` to be very good at reading Excel sheets, and haven't quite efficient with SPSS files.

This package is not doing any person management. We assume that each person has responded to only one test form, once, and we don't check for any duplicates. If users supply a sound person ID of their own, they will be able to link ability measures with results of the same person on a different test, but this is their own responsibility. Also, data from two-stage tests (a routing test and a follow-up test) should be matched by the user before entered.

Any variable whose name has an exact match in the scoring rules input with function `start_new_project` will be treated as an item; any other variables will be treated as person covariates. Any responses to an item that do not have an exact match in the scoring rules will be treated as missing, and ultimately given the lowest score of 0. To score missing data differently, or simply abide to good style, the user can include explicit entries for missing value indicators in the scoring rules.

Value

A list of:

<code>items</code>	The names of the columns in <code>x</code> that were treated as items
<code>covariates</code>	The names of the columns in <code>x</code> that were treated as person covariates
<code>not_listed</code>	A data frame of all responses that will be treated as missing

Author(s)

Ivailo Partchev

`add_item_properties` *Add item properties to a project*

Description

Adds item properties to an existing data base

Usage

```
add_item_properties(db, df)
```

Arguments

<code>db</code>	A handle to the database, i.e. the output of <code>start_new_project</code> or <code>open_project</code>
<code>df</code>	A data frame containing the item properties. See details.

Details

When entering response data in the form of a rectangular person x item table, it is easy to provide person properties but practically impossible to provide item properties. This function provides a possibility to do so. The order of the rows and columns in the data frame is not important but (i) there must be a column called exactly `item` containing the item IDs exactly as entered before, and (ii) all items in the data frame must be known to the data base and all items in the data base must be given properties – otherwise, there will be a warning message, and nothing else will be done. If all is well, the data frame will be added to the data base as table `item_properties`, and any variables in it may be used in analyses involving item properties.

Value

A list of:

unknown_items	Item IDs for any items that were provided in the data frame but could not be found in the data base
items_unaccounted_for	Item IDs for any items that exist in the data base but were not given properties in the data frame

Author(s)

Ivailo Partchev

add_tests	<i>Add test data for an existing booklet</i>
-----------	--

Description

Adds item response data for a test form (a.k.a. booklet) that already exists in the data base. This function has been added as a workaround by user request – we do not necessarily approve.

Usage

```
add_tests(db, x, booklet_id)
```

Arguments

db	A handle to the database, i.e. the output of <code>start_new_project</code> or <code>open_project</code>
x	A data frame containing the responses and, possibly, some additional person characteristics. See details.
booklet_id	A (short) string identifying the test form (booklet)

Value

A list of:

items	The names of the columns in x that were treated as items
covariates	The names of the columns in x that were treated as person covariates
not_listed	A data frame of all responses that will be treated as missing

Author(s)

Ivailo Partchev

distractor_plot	<i>Distractor plot</i>
-----------------	------------------------

Description

Produce a diagnostic distractor plot for an item

Usage

```
distractor_plot(db, item, nc = 1, nr = 1)
```

Arguments

db	A handle to the database, i.e. the output of <code>start_new_project</code> or <code>open_project</code>
item	The ID of the item to plot. A separate plot will be produced for each booklet that contains the item, or an error message if the item ID is not known. Each plot contains a non-parametric regression of each possible response on the total score.
nc	An integer between 1 and 3. Number of columns when putting multiple plots on the same page. Default is 1. May be ignored or adjusted if it does not make sense.
nr	An integer between 1 and 3. Number of rows when putting multiple plots on the same page. Default is 1. May be ignored or adjusted if it does not make sense.

Author(s)

Ivailo Partchev

fit_domains	<i>Estimate the Rasch and the Interaction model per domain</i>
-------------	--

Description

Estimate the parameters of the Rasch model and the Interaction model

Usage

```
fit_domains(db, booklets, domains)
```

Arguments

db	A handle to the database, i.e. the output of <code>start_new_project</code> or <code>open_project</code>
booklets	A vector with the numbers of the test forms (booklets) to be included in the analysis. See Details.
domains	The variable name for the item property defining the domains (subtests)

Details

Unlike the Rasch model, the interaction model cannot be computed concurrently for a whole design of test forms. This function fits the Rasch model and the interaction model on a complete rectangular array of responses, with comparison between the two models playing an important role. The rectangular array may be either one booklet or the intersection (common items) of two or more booklets. If the intersection is empty (no common items), the function will exit with an error message.

Please notice that the booklets are specified with their internal numbers in the data base, not with their names. The names and numbers of the booklets can be inspected with function `show_booklets`.

Value

An object of class `imp` holding results for the Rasch model and the interaction model.

Author(s)

Gunter Maris, Timo Bechger, Ivailo Partchev

fit_models

Estimate the Rasch and the Interaction model

Description

Estimate the parameters of the Rasch model and the Interaction model

Usage

```
fit_models(db, booklets, grpVar = NULL, group = NULL)
```

Arguments

<code>db</code>	A handle to the database, i.e. the output of <code>start_new_project</code> or <code>open_project</code>
<code>booklets</code>	A vector with the numbers of the test forms (booklets) to be included in the analysis. See <code>Details</code> .
<code>grpVar</code>	Variable name for subsetting
<code>group</code>	Value(s) for subsetting

Details

Unlike the Rasch model, the interaction model cannot be computed concurrently for a whole design of test forms. This function fits the Rasch model and the interaction model on a complete rectangular array of responses, with comparison between the two models playing an important role. The rectangular array may be either one booklet or the intersection (common items) of two or more booklets. If the intersection is empty (no common items), the function will exit with an error message.

Please notice that the booklets are specified with their internal numbers in the data base, not with their names. The names and numbers of the booklets can be inspected with function `show_booklets`.

Value

An object of class `rim` holding results for the Rasch model and the interaction model.

Author(s)

Gunter Maris, Timo Bechger, Ivailo Partchev

keys_to_rules *Derive scoring rules from keys*

Description

For multiple choice items that will be scored as 0/1, derive the scoring rules from the keys to the correct responses

Usage

```
keys_to_rules(keys)
```

Arguments

`keys` A data frame containing columns `item`, `nOptions`, and `key` (the spelling is important). See details.

Details

This function might be useful in setting up the scoring rules when all items are multiple-choice and scored as 0/1. (Hint: Because the order in which the scoring rules is not important, one can use the function to generate rules for many MC items and then append per hand the rules for a few complex items.)

The input data frame must contain the exact name of each item, the number of options, and the key. If the keys are all integers, it will be assumed that responses are coded as 1 through `nOptions`. If they are all uppercase letters, it is assumed that responses are coded as A,B,C,... All other cases result in an error.

Value

A data frame that can be used as input to `start_new_project`

Author(s)

Ivailo Partchev

open_project	<i>Open an existing project</i>
--------------	---------------------------------

Description

Opens a data base created by function start_new_project

Usage

```
open_project(db_name = "dexter.db")
```

Arguments

db_name	The name of the data base to be opened.
---------	---

Value

A handle to the data base.

Author(s)

Ivailo Partchev

PISA_item_class	<i>Item properties in the PISA 2012 example</i>
-----------------	---

Description

A data set of item properties in the PISA 2012 example (see the help screen for function add_booklet)

Format

A data set with 109 rows and 6 columns.

plot.rim

*A plot method for the interaction model***Description**

Plot the item-total regressions fit by the interaction (or Rasch) model

Usage

```
## S3 method for class 'rim'
plot(x, items = NULL, summate = TRUE, overlay = FALSE,
      nc = 1, nr = 1, curtains = 10, show.observed = FALSE, ...)
```

Arguments

x	An object produced by function fit_models
items	The items to plot (column numbers). If NULL, all items will be plotted
summate	If FALSE, regressions for polytomous items will be shown for each response option separately; default is TRUE.
overlay	If TRUE and more than one item is specified, there will be two plots, one for the Rasch model and the other for the interaction model, with all items overlaid; otherwise, multiple plots with the two models overlaid. Default is FALSE
nc	An integer between 1 and 3. Number of columns when putting multiple plots on the same page. Default is 1. May be ignored or adjusted if it does not make sense.
nr	An integer between 1 and 3. Number of rows when putting multiple plots on the same page. Default is 1. May be ignored or adjusted if it does not make sense.
curtains	100*the tail probability of the sum scores to be shaded. Default is 10. Set to 0 to have no curtains shown at all.
show.observed	If TRUE, the observed proportion correct at each sum score will be shown as dots. Default is FALSE.
...	Any additional plotting parameters

Author(s)

Ivailo Partchev

print.rim	<i>A print method for the interaction model</i>
-----------	---

Description

Print the available items for plots of the Rasch and the interaction models

Usage

```
## S3 method for class 'rim'
print(x, ...)
```

Arguments

x	An object produced by function fit_models
...	Included to stop check from nagging

Author(s)

Ivailo Partchev

profile_plot	<i>Profile plot</i>
--------------	---------------------

Description

Profile plot

Usage

```
profile_plot(db, booklets, domains, grpVar, group = NULL, model = "IM")
```

Arguments

db	A handle to the database, i.e. the output of start_new_project or open_project
booklets	A vector with the numbers of the test forms (booklets) to be included in the analysis. See Details.
domains	The item property defining the domains
grpVar	Variable name for subsetting
group	Value(s) for subsetting; if NULL, all distinct values will be used.
model	"IM" (default) or "RM"

Value

Nothing interesting

Author(s)

Timo Bechger, Gunter Maris, Ivailo Partchev

show_booklets *List booklets in a project*

Description

Show a list of the test forms (booklets) that have been entered in the db so far

Usage

show_booklets(db)

Arguments

db A handle to the database, i.e. the output of start_new_project or open_project

Value

A data frame showing the internal booklet number, the booklet name, the number persons and the number of items

Author(s)

Ivailo Partchev

show_items *List items in a project*

Description

Show a list of all items that have been entered in the db so far by booklet and position in the booklet

Usage

show_items(db)

Arguments

db A handle to the database, i.e. the output of start_new_project or open_project

Value

A data frame (in string format) showing the booklet design: rows are items, columns are booklets, and the numbers in the cells show the position of the item in the booklet (blank if the booklet does not include the item)

Author(s)

Ivailo Partchev

`show_item_properties` *List item properties*

Description

Show a list of the item properties defined in the project (if any)

Usage`show_item_properties(db)`**Arguments**`db` A handle to the database, i.e. the output of `start_new_project` or `open_project`**Value**

Nothing but messages printed to the screen

Author(s)

Ivailo Partchev

`show_person_properties`
List person properties

Description

Show a list of the person properties defined in the project (if any)

Usage`show_person_properties(db)`**Arguments**`db` A handle to the database, i.e. the output of `start_new_project` or `open_project`**Value**

Nothing but messages printed to the screen

Author(s)

Ivailo Partchev

start_new_project	<i>Start a new project</i>
-------------------	----------------------------

Description

Imports a complete set of scoring rules and starts a new project (data base)

Usage

```
start_new_project(rules, db_name = "dexter.db")
```

Arguments

rules	A data frame with columns item, response, and score. The order is not important but spelling is. Any other columns will be ignored.
db_name	A name for the data base that will be created. If the name does not contain a path, the file will be created in the work directory. Any existing file with the same name will be overwritten.

Details

This package only works with closed items: it does not score any open items. The first step to creating a project is to import an exhaustive list of all items and all admissible responses, along with the score that any of the latter will be given. Responses may be numbers or strings, and they must appear exactly as in the data that will be imported. Scores must be integers, and the minimum score for an item must be 0. When inputting data, all responses not on the list will be treated as missing and ultimately scored 0, but it is good style to include the missing responses in the list. importFrom RSQLite dbConnect dbReadTable dbWriteTable dbListFields importFrom RSQLite dbGetQuery dbSendQuery

Value

If the scoring rules pass a sanity check, a handle to the data base. Otherwise, a data frame listing the problems found.

Author(s)

Ivailo Partchev

`tia_tables`*Simple Test-Item Analysis*

Description

Produce tables with simple test-item analysis statistics

Usage

```
tia_tables(db)
```

Arguments

`db` A handle to the database, i.e. the output of `start_new_project` or `open_project`

Details

When an item is contained in more than one booklet, the item statistics are averaged across booklets with the number of persons in each booklets serving as weights.

Value

A list of two data frames, `itemStats` and `testStats`.

Author(s)

Ivailo Partchev

`verbAggrData`*Verbal aggression data*

Description

A data set of self-reported verbal behaviour in different frustrating situations (Vansteelandt, 2000)

Format

A data set with 316 rows and 25 columns.

verbAggrProperties *Item properties in the verbal aggression data*

Description

A data set of item properties related to the verbal aggression data

Format

A data set with 24 rows and 5 columns.

verbAggrRules *Scoring rules for the verbal aggression data*

Description

A set of (trivial) scoring rules for the verbal aggression data set

Format

A data set with 72 rows and 3 columns (item, response, score).

Index

*Topic **datasets**

- PISA_item_class, 8
- verbAggrData, 14
- verbAggrProperties, 15
- verbAggrRules, 15

add_booklet, 2

add_item_properties, 3

add_tests, 4

distractor_plot, 5

fit_domains, 5

fit_models, 6

keys_to_rules, 7

open_project, 8

PISA_item_class, 8

plot.rim, 9

print.rim, 10

profile_plot, 10

show_booklets, 11

show_item_properties, 12

show_items, 11

show_person_properties, 12

start_new_project, 13

tia_tables, 14

verbAggrData, 14

verbAggrProperties, 15

verbAggrRules, 15