dodoR

April 8, 2015

Type Package

Title Generates IRT plots and item parameter CSVs

Version 1.0

Date 2015-03-25

Author Douglas De Rizzo Meneghetti

Maintainer Douglas De Rizzo Meneghetti <douglasrizzo@fei.edu.br>

Depends R (>= 3.0.2), mirt, latticeExtra

Description More about what it does (maybe more than one line)

License GPL

R topics documented:

| dodoR-package | *R functions I used for my master's thesis* |

Description

This is a personal package I created in order to automate the generation of reports, charts and CSV files concerning the use of Item Response Theory in a binary matrix of responses.

Details

Package:	dodoR
Type:	Package
Version:	1.0
Date:	2015-03-30
License:	What license is it under?

Author(s)

Douglas De Rizzo Meneghetti

Maintainer: Douglas De Rizzo Meneghetti <douglasrizzo@fei.edu.br>

See Also

mirt

| ch | *Enem 2012 Humanity Sciences Test* |

Description

Contains 50000 samples of the brazilian National High School Exam. This dataset contains the 45 humanity sciences questions.

Format

A data frame with 50000 observations on the following 45 variables. Rows represent examinees and columns represent items. Each question has a possible answer of 1, 2, 3, 4 or 5, in numerical form.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

For the creation of these R datasets, the original data, containing all 5.7 million observations, was downloaded from the Brazilian Ministry of Education website. A sample of 50000 observations was then taken, composed only of examinees who answered the full test. Note however, that observations from one dataset available in this package does not directly imply that the observations from the other datasets are indeed from the same examinee, that is, the order of observations is randomized.

Source

http://portal.inep.gov.br/basica-levantamentos-acessar

Examples

```
data(ch)
```

---

ch_keys   *Enem 2012 Humanities Sciences Keys*

Description

The keys for the brazilian National High School Exam humanities sciences test.

Format

A data frame with 1 row and 45 cells, each cell containing the key to the corresponding. Values are numerical, ranging from 1 to 5.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(ch_keys)
```

---

cn	*Enem 2012 Natural Sciences Test*

Description

Contains 50000 samples of the brazilian National High School Exam. This dataset contains the 45 natural sciences questions.

Format

A data frame with 50000 observations on the following 45 variables. Rows represent examinees and columns represent items. Each question has a possible answer of 1, 2, 3, 4 or 5, in numerical form.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

For the creation of these R datasets, the original data, containing all 5.7 million observations, was downloaded from the Brazilian Ministry of Education website. A sample of 50000 observations was then taken, composed only of examinees who answered the full test. Note however, that observations from one dataset available in this package does not directly imply that the observations from the other datasets are indeed from the same examinee, that is, the order of observations is randomized.

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(cn)
```

---

cn_keys *Enem 2012 Natural Sciences Keys*

Description

The keys for the brazilian National High School Exam natural sciences test.

Format

A data frame with 1 row and 45 cells, each cell containing the key to the corresponding. Values are numerical, ranging from 1 to 5.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(cn_keys)
```

---

irt.report *Generate full IRT report based on* mirt *package*

Description

This function generates a full report of an IRT run of the mirt package. It plots all possible charts for the test and each of the items; saves item parameters and item fit statistics (like Zhs and X2) into csv files; and can alternatively be used in conjunction with Sweave to generate a PDF report, since it prints the summaries of all relevant variables during a run.

Usage

```
irt.report(answers, out, out.stats = NULL, out.itemplots = NULL,
  out.testplots = NULL, out.densities = NULL, keys = NULL,
  itemtype = "3PL", method = "EM", optimizer = "NR", verbose = F,
  test_score = T, test_info = T, test_SE = T, test_infoSE = T,
  trace = T, info = T, se = T, score = T, infoSE = T, infotrace = T)
```

Arguments

answers	A data.frame containing examinees answers to items. Rows must represent examinees and columns represent items. If the data.frame is in binary format, keys is not required, but distractor analysis is not perfoermed.
out	A directory to which all files will be saved. If the directory does not exist, it will be recursively created.
out.stats	The directory to which CSV files will be saved. If the directory does not exist, it will be recursively created. if out.stats = NULL, it will be a subfolder of out.
out.itemplots	The directory to which item plots will be saved. If the directory does not exist, it will be recursively created. if out.itemplots = NULL, it will be a subfolder of out.
out.testplots	The directory to which test plots will be saved. If the directory does not exist, it will be recursively created. if out.testplots = NULL, it will be a subfolder of out.
out.densities	The directory to which parameter density plots will be saved. If the directory does not exist, it will be recursively created. if out.densities = NULL, it will be a subfolder of out.
keys	A vector containing the keys of the test, that is, the right answers for each item in the correct order.
itemtype	Item types to be used during parameter estimation. This parameter is passed directly to mirt. Default is '3PL', the three-parameter logistic model. For more options, consult the mirt manual.
method	Estimation method. This parameter is passed directly to mirt. Default is 'EM', the Expectation-Maximization algorithm. For more options, consult the mirt manual.
optimizer	Numerical approximation method used for optimizing convergence of estimation. This parameter is passed directly to mirt. Default is 'NR', the Newton-Raphson method. For more options, consult the mirt manual.
verbose	Defines how much output the algorithm must give. This parameter is passed directly to mirt. Default is FALSE. For more options, consult the mirt manual.
test_score	Whether or not test score plot should be generated.
test_info	Whether or not test information plot should be generated.
test_SE	Whether or not test standard error plot should be generated.
test_infoSE	Whether or not a plot containing both test information and standard error should be generated.
trace	Whether or not item trace plots should be generated.
info	Whether or not item information plots should be generated.
se	Whether or not item standard error plots should be generated.
score	Whether or not item score plots should be generated.
infoSE	Whether or not a plot containing both item information and standard error should be generated.
infotrace	Whether or not a plot containing both item information and trace should be generated.

Value

function returns an object of class SingleGroupClass (SingleGroupClass-class), the same one returned by the 'mirt-package' function 'mirt', in case additional processing is desired beyond the generation of plots and CSV files with parameters

Author(s)

Douglas De Rizzo Meneghetti

---

lce	*Enem 2012 Languages and Codes - Spanish Test*

Description

Contains 50000 samples of the brazilian National High School Exam. This dataset contains the 45 languages and codes questions and the last 5 of them are Spanish language questions.

Format

A data frame with 50000 observations on the following 45 variables. Rows represent examinees and columns represent items. Each question has a possible answer of 1, 2, 3, 4 or 5, in numerical form.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

For the creation of these R datasets, the original data, containing all 5.7 million observations, was downloaded from the Brazilian Ministry of Education website. A sample of 50000 observations was then taken, composed only of examinees who answered the full test. Note however, that observations from one dataset available in this package does not directly imply that the observations from the other datasets are indeed from the same examinee, that is, the order of observations is randomized.

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(lce)
```

`lce_keys` *Enem 2012 Languages and Codes - Spanish Keys*

Description

The keys for the brazilian National High School Exam languages and codes - Spanish test.

Format

A data frame with 1 row and 45 cells, each cell containing the key to the corresponding. Values are numerical, ranging from 1 to 5.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

Source

http://portal.inep.gov.br/basica-levantamentos-acessar

Examples

```
data(lce_keys)
```

`lci` *Enem 2012 Languages and Codes - English Test*

Description

Contains 50000 samples of the brazilian National High School Exam. This dataset contains the 45 languages and codes questions and the last 5 of them are English language questions.

Format

A data frame with 50000 observations on the following 45 variables. Rows represent examinees and columns represent items. Each question has a possible answer of 1, 2, 3, 4 or 5, in numerical form.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

For the creation of these R datasets, the original data, containing all 5.7 million observations, was downloaded from the Brazilian Ministry of Education website. A sample of 50000 observations was then taken, composed only of examinees who answered the full test. Note however, that observations from one dataset available in this package does not directly imply that the observations from the other datasets are indeed from the same examinee, that is, the order of observations is randomized.

Source

http://portal.inep.gov.br/basica-levantamentos-acessar

Examples

```
data(lci)
```

---

lci_keys	*Enem 2012 Languages and Codes - English Keys*

Description

The keys for the brazilian National High School Exam languages and codes - english test.

Format

A data frame with 1 row and 45 cells, each cell containing the key to the corresponding. Values are numerical, ranging from 1 to 5.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(lci_keys)
```

---

mt   *Enem 2012 Mathematics Test*

Description

Contains 50000 samples of the brazilian National High School Exam. This dataset contains the 45 mathematics questions.

Format

A data frame with 50000 observations on the following 45 variables. Rows represent examinees and columns represent items. Each question has a possible answer of 1, 2, 3, 4 or 5, in numerical form.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

For the creation of these R datasets, the original data, containing all 5.7 million observations, was downloaded from the Brazilian Ministry of Education website. A sample of 50000 observations was then taken, composed only of examinees who answered the full test. Note however, that observations from one dataset available in this package does not directly imply that the observations from the other datasets are indeed from the same examinee, that is, the order of observations is randomized.

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(mt)
```

---

mt_keys   *Enem 2012 Languages and Codes - Mathematics Keys*

Description

The keys for the brazilian National High School Exam languages and codes - Mathematics test.

Format

A data frame with 1 row and 45 cells, each cell containing the key to the corresponding. Values are numerical, ranging from 1 to 5.

Details

The Brazilian National High School Exam (Exame Nacional do Ensino Medio, Enem) is a facultative summative educational evaluation available for every Brazilian citizen that has finished, or is in the last year of high school. The test allows students to enroll in higher degree educational courses.

The exam is composed of 4 dichotomous tests: Humanities Sciences, Natural Sciences, Languages and Codes and Mathematics. Each test is composed of 45 multiple-choice items. In the year 2012, the test was taken by approximatelly 5.7 million examinees, whose data is freely available (see Source section).

Source

<http://portal.inep.gov.br/basica-levantamentos-acessar>

Examples

```
data(mt_keys)
```

---

normalize.responses   *Normalize Range of Responses in a Matrix*

Description

normalize.responses takes a response matrix and its keys and puts its values between [1; n of keys for current item]

Usage

```
normalize.responses(x, y)
```

Arguments

x	A response matrix.
y	A list containing the item keys.

Details

Although the use of a non-normalized response-matrix does not interfere in the estimation of IRT parameters, sometimes it is desirable to have item alternatives and its corresponding keys in a normalized range, especially when performing distractor analysis with `irt.report`.

So, for example, a list composed of the values [4, 3, 7] would become [2, 1, 3].

Value

A `list` containing the normalized response matrix in the first slot and its corresponding normalized keys in the second slot.

Author(s)

Douglas De Rizzo Meneghetti

---

sample.rows   *Random Samples of Rows*

Description

`sample.rows` takes a sample of the specified size from the rows of x using either with or without replacement

Usage

```
sample.rows(x, size, replace = F, prob = NULL)
```

Arguments

x	Either a vector of one or more elements from which to choose, or a positive integer.
size	a non-negative integer giving the number of items to choose.
replace	Should sampling be with replacement?
prob	A vector of probability weights for obtaining the elements of the vector being sampled.

Details

It basically does the same as `sample`, but returns row samples instead of column samples.

Value

A `data.frame` containing `size` rows extracted of `x`.

Author(s)

Douglas De Rizzo Meneghetti

References

http://stackoverflow.com/questions/8273313/random-rows-in-dataframe-in-r

Examples

```
##Sample 5 random rows form the cars dataset
sample.rows(x = cars, size = 5)
```

Index