

Package ‘mstDIF’

May 9, 2026

Type Package

Title A Collection of DIF Tests for Multistage Tests

Version 0.1.8

Depends R (>= 4.0.0)

Imports expm, Matrix, PP, mirt (>= 1.31), scDIFtest, eRm

Suggests mvtnorm, testthat, knitr, rmarkdown, strucchange

Description A collection of statistical tests for the detection of differential item functioning (DIF) in multistage tests. Methods entail logistic regression, an adaptation of the simultaneous item bias test (SIBTEST), and various score-based tests. The presented tests provide itemwise test for DIF along categorical, ordinal or metric covariates. Methods for uniform and non-uniform DIF effects are available depending on which method is used.

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.3.1

NeedsCompilation no

VignetteBuilder knitr

Author Rudolf Debelak [aut, cre],
Dries Debeer [aut],
Sebastian Appelbaum [ctb],
Mark J. Gierl [ctb]

Maintainer Rudolf Debelak <rudolf.debelak@gmail.com>

Repository CRAN

Date/Publication 2024-08-30 11:10:03 UTC

Contents

bootstrap_sctest	2
log_reg	4

mstDIF	5
mstDIF-Methods	7
mstSIB	8
permutation_sctest	10
toydata	12

Index	13
--------------	-----------

bootstrap_sctest	<i>A score-based DIF test using the parametric bootstrap approach.</i>
------------------	--

Description

bootstrap_sctest computes a score test to detect DIF in multiple item/parameters with respect to multiple person covariates (DIF_covariate). A parametric bootstrap approach is applied to obtain p-values. That is, given the (item and person) parameters, new data sets are sampled to create the distribution of the test statistic under the null hypothesis. The functionality is limited to the 1-, 2-, and 3-parameter logistic models. Only DIF with respect to the a and b parameters is tested for, which correspond to the item discrimination and the item difficulty parameters.

Usage

```
bootstrap_sctest(
  resp,
  theta = NULL,
  a = rep(1, length(b)),
  b,
  c = rep(0, length(b)),
  DIF_covariate = NULL,
  parameters = c("per_item", "ab", "a", "b"),
  item_selection = NULL,
  nSamples = 1000,
  theta_method = c("wle", "mle", "eap", "map"),
  slope_intercept = FALSE,
  statistic = "auto",
  meanCenter = TRUE,
  decorrelate = FALSE,
  impact_groups = rep(1, dim(resp)[1])
)
```

Arguments

resp	A matrix (or data frame) containing the responses, with the items in the columns.
theta	A vector with the true/estimated ability parameters or NULL (the default) which leads to the ability parameters being estimated.
a	A vector of item slopes/item discriminations.
b	A vector of item locations/item difficulties.

<code>c</code>	A vector of pseudo guessing parameters.
<code>DIF_covariate</code>	A list with the person covariate(s) to test for as element(s).
<code>parameters</code>	A character string, either "per_item", "ab", "a", or "b", to specify which parameters should be tested for.
<code>item_selection</code>	A character vector with the column names or an integer vector with the column numbers in the resp, specifying the items for which the test should be computed. When set to NULL (i.t., the default), all the items are tested.
<code>nSamples</code>	An integer value with the number of permutations to be sampled.
<code>theta_method</code>	A character string, either "wle", "mle", "eap", or "map" that specifies the estimator for the ability estimation. Only relevant when theta == NULL.
<code>slope_intercept</code>	A logical value indicating whether the slope-intercept formulation of the 2-/3-PL model should be used.
<code>statistic</code>	A character string, either "auto", "DM", "CvM", "maxLM", "LMuo", "WDMo", or "maxLMo", specifying the test statistic to be used.
<code>meanCenter</code>	A logical value: should the score contributions be mean centered per parameter?
<code>decorrelate</code>	A logical value: should the score contributions be decorrelated?
<code>impact_groups</code>	A vector indicating impact-group membership for each person.

Details

Author: Dries Debeer

Value

A list with four elements:

`statistics` A matrix containing all the test statistics.

`p` A matrix containing the obtained p -values.

`nSamples` The number of samples taken.

`DIF_covariate` A list containing all the covariate(s) used to order the score contributions, as well as the used test statistics.

See Also

[permutation_sctest](#)

Examples

```
data("toydata")
resp <- toydata$resp
group_categ <- toydata$group_categ
it <- toydata$it
discr <- it[,1]
diff <- it[,2]
```

```
bootstrap_sctest(resp = resp, DIF_covariate = group_categ, a = discr, b = diff,
decorrelate = FALSE)
```

log_reg

A logistic regression DIF test for MSTs

Description

This function allows the detection of itemwise DIF for Multistage Tests. It is based on the comparison of three logistic regression models for each item. The first logistic regression model (Model 1) predicts the positiveness of each response solely on the estimated ability parameters. The second logistic regression model (Model 2) predicts the positiveness based on the ability parameters and the membership to the focal and reference group as additive predictor variables. The third model (Model 3) uses the same predictors as Model 2 to predict the positiveness of the responses, but also includes an interaction effect. Three model comparisons are carried out (Models 1/2, Models 1/3, Models 2/3) based on two criteria: The comparison of the Nagelkerke R squared values, and the p-values of a likelihood ratio test.

Usage

```
log_reg(resp, DIF_covariate, theta = NULL)
```

Arguments

resp	A data frame containing the response matrix. Rows correspond to respondents, columns to items.
DIF_covariate	A factor indicating the membership to the reference and focal groups.
theta	A vector of ability estimates for each respondent.

Details

Author: Sebastian Appelbaum, with minor changes by Rudolf Debelak and Dries Debeer

Value

A list with four elements. The first element is the response matrix, the second element is the name of the DIF covariate, and the third element is the name of the test. The fourth element is a data frame where each row corresponds to an item. The columns of this data frame correspond to the following entries:

N The number of responses observed for this item.

overall_chi_sq The chi squared statistic of the likelihood ratio test comparing Model 1 and Model 3.

overall_p_value The p-values of the likelihood ratio test comparing Model 1 and Model 3 as an indicator for the overall DIF effect.

Delta_NagelkerkeR2 The difference of the Nagelkerke R squared values for Model 1 and Model 3.

UDIF_chi_sq The chi squared statistic of the likelihood ratio test comparing Model 1 and Model 2.

UDIF_p_value The p-values of the likelihood ratio test comparing Model 1 and Model 2.

UDIF_Delta_NagelkerkeR2 The difference of the Nagelkerke R squared values for Model 1 and Model 2.

CDIF_chi_sq The chi squared statistic of the likelihood ratio test comparing Model 2 and Model 3.

CDIF_p_value The p-values of the likelihood ratio test comparing Model 2 and Model 3.

CDIF_Delta_NagelkerkeR2 The difference of the Nagelkerke R squared values for Model 2 and Model 3.

Examples

```
data("toydata")
resp <- toydata$resp
group_categ <- toydata$group_categ
theta_est <- toydata$theta_est
log_reg(resp, DIF_covariate = factor(group_categ), theta = theta_est)
```

mstDIF

A general function to detect differential item functioning (DIF) in multistage tests (MSTs)

Description

This function allows the application of various methods for the detection of differential item functioning in multistage tests. Currently five methods are implemented: 1. Logistic Regression, 2. mstSIB, 3. analytical score-base tests, 4. a score-based Bootstrap test, 5. a score-based permutation test. The required input depends on the chosen DIF test.

Usage

```
## Default S3 method:
mstDIF(resp, DIF_covariate, method, theta = NULL, see = NULL, ...)

## S3 method for class 'AllModelClass'
mstDIF(
  object,
  DIF_covariate,
  method,
  theta = NULL,
  see = NULL,
```

```

    theta_method = "WLE",
    ...
)

## S3 method for class 'dRm'
mstDIF(object, DIF_covariate, method, theta = NULL, see = NULL, ...)

```

Arguments

<code>resp, object</code>	A data frame or matrix containing the response matrix. Rows correspond to respondents, columns to items. Or an object of class <code>SingleGroup</code> -class or <code>MultiGroup</code> -class object as returned by <code>mirt</code> , or a <code>dRm</code> object as returned by the <code>RM</code> function in <code>eRm</code> .
<code>DIF_covariate</code>	A vector of ability estimates for each respondent.
<code>method</code>	A character value indicating the DIF test that should be used. Possible values are "logreg" (Logistic regression), "mstsib" (mstSIB), "bootstrap" (score-based Bootstrap test), "permutation" (score-based) permutation test) and "analytical" (analytical score-based test).
<code>theta</code>	Estimates of the ability parameters.
<code>see</code>	Estimates of the standard error of estimation.
<code>...</code>	Additional, test-specific arguments.
<code>theta_method</code>	Method for estimating the ability parameters if they should be estimated based on the responses. The calculation is carried out by the <code>mirt</code> package. Can be: "WLE" (default), "MAP", "EAP", "ML", "EAPsum", "plausible", "classify".

Details

Author: Rudolf Debelak and Dries Debeer

Value

An object of class `mstDIF`, which is a list with the following elements:

<code>resp</code>	The response matrix as a data frame.
<code>method</code>	The used DIF detection method.
<code>test</code>	The used test or statistic.
<code>DIF_covariate</code>	The person covariate tested for DIF.
<code>DIF_test</code>	A list with the DIF-test results.
<code>call</code>	The function call.
<code>method_results</code>	The complete output of the selected DIF test. Details depend on the DIF test.

Methods (by class)

- `mstDIF(default)`: Default `mstDIF` method
- `mstDIF(AllModelClass)`: `mstDIF` method for `mirt`-objects
- `mstDIF(dRm)`: `mstDIF` method for `dRm`-objects

See Also[mstDIF-Methods](#)**Examples**

```
## load data
data("toydata")
resp <- toydata$resp
group_categ <- factor(toydata$group_categ)
theta_est <- toydata$theta_est
see_est <- toydata$see_est

## test DIF along a categorical covariate (a factor) using the
## logistic regression method
res1 <- mstDIF(resp, DIF_covariate = group_categ, method = "logreg",
theta = theta_est)
res1
summary(res1)

## test DIF along a categorical covariate (a factor) using the
## mstSIB method
res2 <- mstDIF(resp, DIF_covariate = factor(group_categ), method = "mstsib",
theta = theta_est, see = see_est)
res2
summary(res2)
```

mstDIF-Methods

Methods for the mstDIF-class

Description

print and summary methods for objects of the mstDIF-class, as returned by [mstDIF](#). See details for more information about the methods.

Usage

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

## S3 method for class 'mstDIF'
summary(object, DIF_type = "overall", ordered = TRUE, ...)
```

Arguments

x	an object of class mstDIF
...	other arguments passed to the method.
object	an object of class mstDIF

DIF_type	a string that should one or more of "overall", "uniform", "non-uniform", "all".
ordered	logical: should the summary be ordered according to the obtained p-values (in ascending order)?

Details

The print method prints some basic information about the mstDIF-class object.

The summary method computes a data frame with a row for each item that was included in the test. The columns are:

item The name of the item

statistic The value for the used statistic per item

p_value The p-value per item

eff_size An effect-size for the DIF-test, if applicable

Examples

```
## load data
data("toydata")

## fit 2PL model using mirt
mirt_model <- mirt::mirt(toydata$resp, model = 1)

## test DIF along a continuous covariate
DIFtest <- mstDIF(mirt_model, DIF_covariate = toydata$group_cont,
method = "analytical")

## print
DIFtest

## summary
summary(DIFtest)
```

mstSIB

The mstSIB test for MSTs

Description

This function allows the detection of itemwise DIF using the mstSIB test.

Usage

```
mstSIB(
  resp,
  DIF_covariate,
  theta = NULL,
  see = NULL,
  cellmin = 3,
  pctmin = 0.9,
  NCell = 80
)
```

Arguments

<code>resp</code>	A data frame containing the response matrix. Rows correspond to respondents, columns to items.
<code>DIF_covariate</code>	A vector indicating the membership to the reference (0) and focal (1) groups.
<code>theta</code>	A vector of ability estimates for each respondent.
<code>see</code>	A vector of the standard error of the ability estimates for each respondent.
<code>cellmin</code>	Minimum number of respondents per cell for the focal and reference group. Cells with fewer respondents are discarded.
<code>pctmin</code>	Minimum rate of focal and reference group that should be used for estimating the over ability difference between focal and groups after discarding cells with few respondents.
<code>NCell</code>	The initial number of cells for estimating the overall ability difference between the focal and reference groups.

Details

Author: Mark J. Gierl, with minor changes by Rudolf Debelak and Dries Debeer

Value

A list with four elements. The first element is the response matrix, the second element is the name of the DIF covariate, and the third element is the name of the test. The fourth element is a matrix where each row corresponds to an item. The columns correspond to the following entries:

- Beta** The estimated weighted ability difference between the focal and reference groups.
- Vars** The estimation error of the weighted ability difference between the focal and reference groups.
- N_R** The number of respondents in the reference group.
- N_F** The number of respondents in the focal group.
- NCell** The initial number of cells for estimating the overall ability difference between the focal and reference groups.
- p_value** The p-value of the null hypothesis that the ability difference between the focal and reference groups is 0.

Examples

```
data("toydata")
resp <- toydata$resp
group_categ <- toydata$group_categ
theta_est <- toydata$theta_est
see_est <- toydata$see_est
mstSIB(resp = as.data.frame(resp), theta = theta_est,
DIF_covariate = group_categ, see = see_est)
```

permutation_sctest *A score-based DIF test using the permutation approach.*

Description

permutation_sctest computes a score test to detect DIF in multiple item/parameters with respect to multiple person covariates (DIF_covariate). A resampling approach is applied to obtain p-values. That is, given the (item and person) parameters, new data sets are sampled to create the distribution of the test statistic under the null hypothesis. The functionality is limited to the 1-, 2-, and 3-parameter logistic models. Only DIF with respect to the a and b parameters is tested for, which correspond to the item discrimination and the item difficulty parameters.

Usage

```
permutation_sctest(
  resp,
  theta = NULL,
  a = rep(1, length(b)),
  b,
  c = rep(0, length(b)),
  DIF_covariate = NULL,
  parameters = c("per_item", "ab", "a", "b"),
  item_selection = NULL,
  nSamples = 1000,
  theta_method = c("wle", "mle", "eap", "map"),
  slope_intercept = FALSE,
  statistic = "auto",
  meanCenter = TRUE,
  decorrelate = FALSE,
  impact_groups = rep(1, dim(resp)[1])
)
```

Arguments

resp	A matrix (or data frame) containing the responses, with the items in the columns.
theta	A vector with the true/estimated ability parameters or NULL (the default) which leads to the ability parameters being estimated.

a	A vector of item slopes/item discriminations.
b	A vector of item locations/item difficulties.
c	A vector of pseudo guessing parameters.
DIF_covariate	A list with the person covariate(s) to test for as element(s).
parameters	A character string, either "per_item", "ab", "a", or "b", to specify which parameters should be tested for.
item_selection	A character vector with the column names or an integer vector with the column numbers in the resp, specifying the items for which the test should be computed. When set to NULL (i.t., the default), all the items are tested.
nSamples	An integer value with the number of permutations to be sampled.
theta_method	A character string, either "wle", "mle", "eap", or "map" that specifies the estimator for the ability estimation. Only relevant when theta == NULL.
slope_intercept	A logical value indicating whether the slope-intercept formulation of the 2-/3-PL model should be used.
statistic	A character string, either "auto", "DM", "CvM", "maxLM", "LMuo", "WDMo", or "maxLMo", specifying the test statistic to be used.
meanCenter	A logical value: should the score contributions be mean centered per parameter?
decorrelate	A logical value: should the score contributions be decorrelated?
impact_groups	A vector indicating impact-group membership for each person.

Details

Author: Dries Debeer

Value

A list with four elements:

`statistics` A matrix containing all the test statistics.

`p` A matrix containing the obtained *p*-values.

`nSamples` The number of samples taken.

`DIF_covariate` A list containing all the covariate(s) used to order the score contributions, as well as the used test statistics.

See Also

[bootstrap_sctest](#)

Examples

```
data("toydata")
resp <- toydata$resp
group_categ <- toydata$group_categ
it <- toydata$it
discr <- it[,1]
```

```
diff <- it[,2]

permutation_sctest(resp = resp, DIF_covariate = group_categ, a = discr, b = diff,
decorrelate = FALSE)
```

toydata

A Toy Example of 1000 Respondents Working on a Multistage Test

Description

Data of 1000 respondents working on a multistage test using a (1,2,2) design. The responses were generated based on the 2PL model. Each module consists of 7 items. Data were generated using the mstR package, version 1.2 (<https://cran.r-project.org/web/packages/mstR/index.html>).

Usage

```
toydata
```

Format

A list with 7 elements:

resp The response matrix, with rows corresponding to respondents and columns corresponding to items.

it A matrix of item parameters. The columns contain the discrimination, difficulty, pseudo-guessing and inattention parameters of the 4PL model. The discrimination parameters were drawn from a $N(1,0.2)$ distribution. The difficulty parameters were drawn from normal distributions. For module 1 (items 1-7), this distributions was $N(0,1)$, for modules 2 and 4 (items 8-14 and 22-28) it was $N(1,1)$ and for modules 3 and 5 (items 15-21 and 29-35) the distribution was $N(-1,1)$.

theta The true ability parameters.

theta_est The ability parameters estimated by the WLE estimator.

group_categ A simulated categorical person covariate. The first 500 respondents belong to group 0, the remaining 500 respondents to group 1.

group_cont A simulated continuous person covariate. It simulates an age covariate, with a uniform distribution between 20 and 60.

see_est The standard errors of the estimated ability parameters.

Index

* datasets

toydata, [12](#)

bootstrap_sctest, [2](#), [11](#)

log_reg, [4](#)

mstDIF, [5](#), [7](#)

mstDIF-Methods, [7](#)

mstSIB, [8](#)

permutation_sctest, [3](#), [10](#)

print.mstDIF (mstDIF-Methods), [7](#)

summary.mstDIF (mstDIF-Methods), [7](#)

toydata, [12](#)