Package ‘DIFtree’

October 12, 2022

Type Package

Title Item Focussed Trees for the Identification of Items in Differential Item Functioning

Version 3.1.6

Date 2020-06-04

Author Moritz Berger

Depends penalized, VGAM, grid, gridBase, plotrix

Maintainer Moritz Berger <moritz.berger@imbie.uni-bonn.de>

Description Item focussed recursive partitioning for simultaneous selection of items and variables that induce Differential Item Functioning (DIF) in dichotomous or polytomous items.

License GPL-2

LazyLoad yes

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-05 09:30:03 UTC

R topics documented:

data_sim_PCM .............................................................. 2
data_sim_Rasch .......................................................... 3
DIFtree ................................................................. 4
plot.DIFtree ........................................................... 6
predict.DIFtree ....................................................... 8
summary.DIFtree ..................................................... 9

Index 11
**data_sim_PCM**

Simulated Data Set with Polytomous Items

**Description**

The data set is simulated from a Partial Credit Model where some items exhibit differential item functioning. Existing differences in item difficulties are simulated by step-functions. The true, simulated DIF structure is described in Bollmann et al. (2017), Section 4.3.

**Usage**

```r
data(data_sim_PCM)
```

**Format**

A data frame containing 500 observations on 4 variables:

- **Y** matrix with categorical responses (3-point scale)
- **x1** binary covariate
- **x2** ordinal covariate
- **x3** numeric covariate

**References**


**Examples**

```r
data(data_sim_PCM)

Y <- data_sim_PCM[,1]
X <- data_sim_PCM[,,-1]

apply(Y,2,table)
summary(X)
```
Simulation Data Set with Dichotomous Items

Description

The data set is simulated from a Rasch model where some items exhibit differential item functioning. Existing differences in item difficulties are simulated by step-functions. The true, simulated DIF structure is described in Tutz and Berger (2015), Section 4.2.

Usage

data(data_sim_Rasch)

Format

A data frame containing 500 observations on 5 variables:

Y matrix with binary 0/1 response for 20 items
x1 binary covariate 1
x2 metric covariate 1
x3 binary covariate 2
x4 metric covariate 2

References


Examples

data(data_sim_Rasch)

Y <- data_sim_Rasch[,1]
X <- data_sim_Rasch[,1]

hist(rowSums(Y), breaks = 0:19 + 0.5)
supply(X)
DIFtree

Item focussed Trees for the Identification of Items in Differential Item Functioning

Description
A function to estimate item focussed trees for simultaneous selection of items and variables that induce DIF (Differential Item Functioning) in dichotomous or polytomous items. DIF detection can be based on the Rasch Model (dichotomous case), the Logistic Regression Approach (dichotomous case) or the Partial Credit Model (polytomous case). The basic method of item focussed recursive partitioning in Rasch Models is described in Tutz and Berger (2015).

Usage
DIFtree(Y, X, model = c("Rasch", "Logistic", "PCM"), type = c("udif", "dif", "nudif"), alpha = 0.05, nperm = 1000, trace = FALSE, penalize = FALSE, ...)

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

Arguments
Y
Matrix or Data.frame of binary 0/1 or categorical response (rows correspond to persons, columns correspond to items)

X
Data.frame of (not scaled) covariates (rows correspond to persons, columns correspond to covariates)

model
Type of model to be fitted; can be "Rasch", "Logistic" or "PCM".

type
Type of DIF to be modelled; one out of "udif", "dif" and "nudif". For "Rasch" and "PCM" only uniform DIF can be modelled and therefore type will be ignored.

alpha
Global significance level for the permutation tests

nperm
Number of permutations used for the permutation tests

trace
If true, information about the estimation progress is printed

penalize
If true, a small ridge penalty is added to ensure existence of model parameters; only for "Rasch".

...
Further arguments passed to or from other methods

x
Object of class "DIFtree"

Details
The methods require 0/1 coded answers on binary items ("Rasch" and "Logistic") or categorical answers on polytomous items ("PCM"). Items with DIF are gradually identified by recursive partitioning.
For "Rasch" one yields a model with linear predictors

\[ \eta_{pi} = \theta_p - \tau_i(x_p), \]

where \( \theta_p \) correspond to the ability and \( x_p \) correspond to the covariate vector of person \( p \).

For "Logistic" one yields a model with linear predictors

- **Uniform DIF**, type="udif"
  \[ \eta_{pi} = S_p \beta_i + \tau_i(x_p), \]
  where \( S_p \) corresponds to the test score and \( x_p \) corresponds to the covariate vector of person \( p \).

- **DIF and Non-Uniform DIF**, type="dif", "nudif"
  \[ \eta_{pi} = \tau_i(x_p) + \tau_i(S_p, x_p), \]
  where \( S_p \) corresponds to the test score and \( x_p \) corresponds to the covariate vector of person \( p \).

For "PCM" one yields a model with linear predictors

\[ \eta_{pir} = \theta_p - \tau_{ir}(x_p), \]

where \( \theta_p \) correspond to the ability and \( x_p \) correspond to the covariate vector of person \( p \).

Significance of each split is verified by permutation tests. The result of the permutation tests can strongly depend on the number of permutations \( n_{perm} \). In the case of pure terminal nodes estimates of the model do not exist. If \( \text{penalize}=\text{TRUE} \) a small ridge penalty is added during estimation to ensure existence of all parameters.

**Value**

Object of class "DIFtree". An object of class "DIFtree" is a list containing the following components:

- **splits** Matrix with detailed information about all executed splits during the estimation process
- **coefficients** List of estimated coefficients for items with and without DIF. Structure of coefficients depends on model and type.
- **pvalues** P-values of each permutation test during the estimation process
- **devs** Maximal value statistics \( T_j \) of the selected variables in each iteration during the estimation process
- **crit** Critical values of each permutation test during the estimation process
- **Y** Response matrix used in the estimation process
- **X** Model matrix used in the estimation process
- **persons** Number of persons
- **items** Number of items

**Author(s)**

Moritz Berger <moritz.berger@imbie.uni-bonn.de>
http://www.imbie.uni-bonn.de/personen/dr-moritz-berger/
References


See Also

plot.DIFtree, predict.DIFtree, summary.DIFtree

Examples

data(data_sim_Rasch)
data(data_sim_PCM)

Y1 <- data_sim_Rasch[,1]
X1 <- data_sim_Rasch[-1]

Y2 <- data_sim_PCM[,1]
X2 <- data_sim_PCM[-1]

## Not run:
mod1 <- DIFtree(Y=Y1,X=X1,model="Logistic",type="udif",alpha=0.05,nperm=1000,trace=TRUE)
print(mod1)

mod2 <- DIFtree(Y=Y2,X=X2,model="PCM",alpha=0.05,nperm=100,trace=TRUE)
print(mod2)

## End(Not run)

plot.DIFtree

Plotting of Item focussed Trees

Description

Visualization of trees for items with DIF identified by item focussed recursive partitioning in dichotomous or polytomous items.
## S3 method for class 'DIFtree'

```r
plot(x, item, component = "intercept", cex.lines = 2,
    cex.branches = 1, cex.coefs = 1, cex.main = 1, title = NULL, ...)
```

### Arguments

- `x`: Object of class `DIFtree`
- `item`: Number of the item, for which the tree shall be plotted
- `component`: Component of the model for which the tree shall be plotted; can be "intercept" or "slope". For "Rasch" and "PCM" only one tree of item parameters is available for each DIF item and therefore component will be ignored.
- `cex.lines`: Width of branches of the tree
- `cex.branches`: Size of the labels of branches of the tree
- `cex.coefs`: Size of coefficients in the terminal nodes of the tree
- `cex.main`: Size of the title of the tree
- `title`: Optional title, which is added to the tree; if `title=NULL` the title is the number of the plotted item.
- `...`: Further arguments passed to or from other methods

### Author(s)

Moritz Berger <moritz.berger@imbie.uni-bonn.de>

[http://www.imbie.uni-bonn.de/personen/dr-moritz-berger/](http://www.imbie.uni-bonn.de/personen/dr-moritz-berger/)

### References


### See Also

`DIFtree`, `predict.DIFtree`, `summary.DIFtree`

### Examples

```r
data(data_sim_Rasch)
Y <- data_sim_Rasch[,1]
X <- data_sim_Rasch[,1]
```
## Not run:

mod <- DIFtree(Y=Y,X=X,model="Logistic",type="udif",alpha=0.05,nperm=1000,trace=TRUE)
plot(mod,item=1)
## End(Not run)

### Description

The function returns predictions of item parameters obtained by item focussed recursive partitioning in dichotomous or polytomous items.

### Usage

```r
## S3 method for class 'DIFtree'
predict(object, item, newdata, ...)
```

### Arguments

- **object**: Object of class `DIFtree`
- **item**: Number of the item, for which the prediction shall be returned
- **newdata**: New data.frame, for which the prediction shall be returned
- **...**: Further arguments passed to or from other methods

### Details

For "Rasch" model the function returns the predicted item difficulty. For "Logistic" models the function returns the predicted intercept and/or slope. For "PCM" the function returns the predicted threshold parameters.

### Author(s)

Moritz Berger <moritz.berger@imbie.uni-bonn.de>

http://www.imbie.uni-bonn.de/personen/dr-moritz-berger/

### References


See Also

DIFtree, plot.DIFtree, summary.DIFtree

Examples

```r
data(data_sim_Rasch)
Y <- data_sim_Rasch[,1]
X <- data_sim_Rasch[,-1]
Xnew <- data.frame("x1"=c(0,1),"x2"=c(-1.1,2.5),"x3"=c(1,0),"x4"=c(-0.2,0.7))

## Not run:
mod <- DIFtree(Y=Y,X=X,model="Logistic",type="udif",alpha=0.05,nperm=1000,trace=TRUE)
predict(mod,item=1,Xnew)
## End(Not run)
```

summary.DIFtree

Summary for fitted Item focussed Trees

Description

The function takes an object of class "DIFtree" and returns an useful summary with an overview of all executed splits during the estimation procedure.

Usage

```r
## S3 method for class 'DIFtree'
summary(object, ...)

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

Arguments

- **object**: Object of class **DIFtree**
- **...**: Further arguments passed to or from other methods
- **x**: Object of class **summary.DIFtree**
summary.DIFtree

Value

Object of class "summary.DIFtree". An object of class "summary.DIFtree" is a list containing the following components:

- **stats**: Useful overview of detected DIF items, responsible variables and executed splits
- **nosplits**: Total number of executed splits during the estimation procedure

Author(s)

Moritz Berger <moritz.berger@imbie.uni-bonn.de>
http://www.imbie.uni-bonn.de/personen/dr-moritz-berger/

References


See Also

DIFtree, plot.DIFtree, predict.DIFtree

Examples

```r
data(data_sim_Rasch)

Y <- data_sim_Rasch[,1]
X <- data_sim_Rasch[,-1]

## Not run:
mod <- DIFtree(Y=Y,X=X,model="Logistic",type="udif",alpha=0.05,nperm=1000,trace=TRUE)
summary(mod)

## End(Not run)
```
Index

data_sim_PCM, 2
data_sim_Rasch, 3
DIFtree, 4, 7–10

plot.DIFtree, 6, 6, 9, 10
predict.DIFtree, 6, 7, 8, 10
print.DIFtree(DIFtree), 4
print.summary.DIFtree
  (summary.DIFtree), 9

summary.DIFtree, 6, 7, 9, 9