

# Package ‘phd’

October 14, 2022

**Type** Package

**Title** Permutation Testing in High-Dimensional Linear Models

**Version** 0.2

**Date** 2022-01-05

**Author** Jesse Hemerik, Livio Finos

**Maintainer** Jesse Hemerik <jesse.hemerik@wur.nl>

**Description** Provides permutation methods for testing in high-dimensional linear models. The tests are often robust against heteroscedasticity and non-normality and usually perform well under anti-sparsity. See Hemerik, Thoresen and Finos (2021) <[doi:10.1080/00949655.2020.1836183](https://doi.org/10.1080/00949655.2020.1836183)>.

**License** GNU General Public License

**Imports** methods, stats, glmnet

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2022-01-06 00:20:05 UTC

## R topics documented:

|              |          |
|--------------|----------|
| doublers     | 1        |
| FLhd         | 2        |
| <b>Index</b> | <b>4</b> |

---

|          |                                                         |
|----------|---------------------------------------------------------|
| doublers | <i>Permutation test based on double residualization</i> |
|----------|---------------------------------------------------------|

---

## Description

Provides a class of tests for testing in high-dimensional linear models. The tests are robust against heteroscedasticity and non-normality. They often provide good type I error control even under anti-sparsity.

**Usage**

```
doublerses(y,X,X1,nperm=2E4,lambda="lambda.min",flip="FALSE",nfolds=10)
```

**Arguments**

|        |                                                                                                                                                                                                                                                                                        |
|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| y      | The values of the outcome.                                                                                                                                                                                                                                                             |
| X      | The design matrix. If the covariate of interest is included in X, it should be included in the first column. If it is not included in X, then specify X1. The data do not need to be standardized, since this is automatically done by this function. Do not include a columns of 1's. |
| X1     | n-vector with the (1-dimensional) covariate of interest. X1 should only be specified if the covariate of interest is not already included in X.                                                                                                                                        |
| nperm  | The number of random permutations (or sign-flipping maps) used by the test                                                                                                                                                                                                             |
| lambda | The penalty used in the ridge regressions. Default is "lambda.min", which means that the penalty is obtained using cross-validation. One can also enter "lambda.1se", which is an upward-conservative estimate of the optimal lambda.                                                  |
| flip   | Default is "FALSE", which means that permutation is used. If "TRUE", then sign-flipping is used.                                                                                                                                                                                       |
| nfolds | The number of folds used in the cross-validation (in case lambda is determined using cross-validation).                                                                                                                                                                                |

**Value**

A two-sided p-value.

**Examples**

```
set.seed(5193)
n=30

X <- matrix(nr=n,nc=60,rnorm(n*60))
y <- X[,1]+X[,2]+X[,3] + rnorm(n,mean=0) #H0: first coefficient=0. So H0 is false

doublerses(y, X, nperm=2000, lambda=100,flip="FALSE")
```

**Description**

Provides a class of tests for testing in high-dimensional linear models. The tests are robust against heteroscedasticity and non-normality. They often provide good type I error control even under anti-sparsity.

**Usage**

```
FLhd(y,X,X1,nperm=2E4,lambda="lambda.min",flip="FALSE",nfolds=10,statistic="partialcor")
```

**Arguments**

|           |                                                                                                                                                                                                                                                                                        |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| y         | The values of the outcome.                                                                                                                                                                                                                                                             |
| X         | The design matrix. If the covariate of interest is included in X, it should be included in the first column. If it is not included in X, then specify X1. The data do not need to be standardized, since this is automatically done by this function. Do not include a columns of 1's. |
| X1        | n-vector with the (1-dimensional) covariate of interest. X1 should only be specified if the covariate of interest is not already included in X.                                                                                                                                        |
| nperm     | The number of random permutations (or sign-flipping maps) used by the test                                                                                                                                                                                                             |
| lambda    | The penalty used in the ridge regressions. Default is "lambda.min", which means that the penalty is obtained using cross-validation. One can also enter "lambda.1se", which is an upward-conservative estimate of the optimal lambda.                                                  |
| flip      | Default is "FALSE", which means that permutation is used. If "TRUE", then sign-flipping is used.                                                                                                                                                                                       |
| statistic | The type of statistic that is used within the permutation test. Either the partial correlation ("partialcor") or the semi-partial correlation ("semipartialcor").                                                                                                                      |
| nfolds    | The number of folds used in the cross-validation (in case lambda is determined using cross-validation).                                                                                                                                                                                |

**Value**

A two-sided p-value.

**Examples**

```
set.seed(5193)
n=30

X <- matrix(nr=n,nc=60,rnorm(n*60))
y <- X[,1]+X[,2]+X[,3] + rnorm(n,mean=0) #H0: first coefficient=0. So H0 is false

FLhd(y, X, nperm=2000, lambda=100,flip="FALSE", statistic="partialcor")
```

# Index

doubles, 1

FLhd, 2