Package 'PIGE'

August 30, 2017

Type Package

Title Self Contained Gene Set Analysis for Gene- And Pathway-Environment Interaction Analysis

Version 1.1

Date 2017-08-29

Author Benoit Liquet, Therese Truong, Camilo Broc

Maintainer Benoit Liquet <benoit.liquet@univ-pau.fr>

Depends snowfall,ARTP,xtable,survival

Description

Extension of the 'ARTP' package (Yu, K., et al. (2009) <doi:10.1002/gepi.20422>) for gene- and pathway-environment interaction. A permutation and a parametric bootstrap approaches have been implemented for investigating gene- and pathway-environment interaction analysis (Truong, T., et al. (2014) <doi:10.1530/ERC-14-0121>).

License GPL (>= 2.0)

NeedsCompilation no

Repository CRAN

Date/Publication 2017-08-30 08:23:09 UTC

R topics documented:

`PIGE-package` *Gene and pathway p-values using the Adaptive Rank Truncated Product test*

Description

An R package for computing gene and pathway p-values using the Adaptive Rank Truncated test (ARTP). This package can be used to analyze pathways/genes based on a genetic association study, with a binary case-control outcome or a survival outcome. This package is an extension of the ARTP method developped by Kai Yu (Genet Epidemiol. 2009) for gene- and pathway-environment interaction analysis.

Details

The statistical significance of the pathway-level test statistics is evaluated using a highly efficient permutation algorithm that remains computationally feasible irrespective of the size of the pathway and complexity of the underlying test statistics for summarizing SNP- and gene-level associations. The function `ARTP.GE` is used to compute gene and pathway p-values provided that the observed and permutation p-values for each SNP already exist in files. The input files required for `ARTP.GE` could be obtained by calling the function `permutation.snp` and the function `compute.p.snp.obs`.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

References

Yu K, Li Q, Berger AW, Pfeiffer R, Rosenberg P, Caporaso N, Kraft P, Chatterjee N (2009). Pathway analysis by adaptive combination of P-values. Genet Epidemiol 33:700-709.

See Also

`ARTP.GE`, `permutation.snp`, `compute.p.snp.obs`

ARTP.GE *Gene and pathway p-values using ARTP method*

Description

Calculate gene- and pathway-environment interaction p-values using the Adaptive Rank Truncated Product method. This function uses mainly the function ARTP_pathway developped by Kai Yu (R package ARTP).

Usage

```
ARTP.GE(data.gene.pathway, list.gene.snp, p.snp.permut, p.snp.obs,
  inspect.snp.n = 1, inspect.snp.percent = 0, inspect.gene.n = 10,
  inspect.gene.percent = 0.05, temp.dir = "TEMP/", nperm)
```

Arguments

`data.gene.pathway`
: Data frame (Gene X Pathways) of 0 and 1 values. The rownames (gene name considered) and the colnames (names of the studied pathways) have to be specified. The value 1 indicates that a gene is included in the corresponding pathway.

`list.gene.snp`
: List containing for each gene the corresponding SNP ids. This list could be generated by `data.to.list.gene.snp` function.

`p.snp.permut`
: the output matrix from either `permutation.snp`, `bootstrap.snp` or a file with the SNP ids and p-values (see details).

`p.snp.obs`
: The output data frame from `compute.p.snp.obs` or a file with the SNP ids and p-values (see details).

`inspect.snp.n`
: The number of candidate truncation points to inspect the top SNPs in a gene. The default is 1.

`inspect.snp.percent`
: A value x between 0 and 1 such that a truncation point will be defined at every x percent of the top SNPs. The default is 0 so that the truncation points will be `1:inspect.snp.n`

`inspect.gene.n`
: The number of candidate truncation points to inspect the top genes in the pathway. The default is 10.

`inspect.gene.percent`
: A value x between 0 and 1 such that a truncation point will be defined at every x percent of the top genes. The default is 0.05.

`temp.dir`
: A folder to keep temporary files that will be created.

`nperm`
: Number of permutation used or number of parametric bootsrap used.

Details

If the p-values are not computed using permutation.snp, bootstrap.snp and compute.p.snp.obs then the format for p.snp.obs and p.snp.permut should be as follows. Both files must be uncompressed, comma seperated files with the first row as the SNP ids in the same order. Row 2 of obs.file has the observed p-values, and starting from row 2 in perm.file are the permuted p-values or the boostrap p-values. A random seed should be set before calling ARTP.GE.R in order to reproduce results. The randomness is due to the ranking of p-values, where ties are broken randomly.

Value

The returned value is a list with names "res.gene.list" and "res.pathway". res.gene.list is a list with length equals to the number of investigated pathways. Each element of the list is a data frame containing the gene name, number of SNPs belonging to the gene that were included in the analysis, and the ARTP p-value for the gene. res.pathway contains the ARTP p-values for all the pathway analysed. The results contained in res.pathway are saved in a file named "ARTP-GEI.RData". The data frame containing all the gene analysed, the number of SNPs belonging to each gene and the ARTP p-value are saved in a file named "ARTP-GENE.RData".

Author(s)

Benoit Liquet <benoit.liquet@univ-pau.fr>
Therese Truong <therese.truong@inserm.fr>

References

Yu K, Li Q, Berger AW, Pfeiffer R, Rosenberg P, Caporaso N, Kraft P, Chatterjee N (2009). Pathway analysis by adaptive combination of P-values. Genet Epidemiol 33:700-709.

Examples

```
data(data.pathway)
data(list.gene.snp)
## Not run:
data(data.pige)

###First example: compute observed p-value (orignal data) and permuted p-value
res <-data.to.PIGE(data=data.pige,data.pathway=data.pathway,
list.gene.snp=list.gene.snp,choice.pathway=c(1,2))
formul <- formula(y~factor(cov1)+factor(cov2)+factor(cov3)+factor(cov4)
+var_int)
p.snp.obs.ex <-  compute.p.snp.obs(data=data.pige,model=formul,
indice.snp=res$snp.selected,var.inter="var_int",class.inter=NULL)
p.snp.permut.ex <-  permutation.snp(model=formul,data=data.pige,
indice.snp=res$snp.selected,var.inter="var_int",class.inter=NULL,
nbcpu=3,Npermut=9,file.out="res-permut")
set.seed(10)
result.1 <- ARTP.GE(data.gene.pathway=data.pathway,
list.gene.snp=list.gene.snp,p.snp.permut=p.snp.permut.ex,
p.snp.obs=p.snp.obs.ex,inspect.snp.n=5,inspect.snp.percent=0.05
,inspect.gene.n=10,inspect.gene.percent=0.05,temp.dir="TEMP/"
,nperm=9)
```

```
result.1

##Second example: observed and permuted p-values have already been computed
path.data <- paste(system.file("sampleData", package="PIGE"),"/",sep="")
res.permut <- read.table(file=paste(path.data,"res-permut.txt",sep="")
,header=TRUE,sep=" ")
res.obs   <- read.table(file=paste(path.data,"res-obs.txt",sep="")
,header=TRUE,sep=" ")
result.2 <- ARTP.GE(data.gene.pathway=data.pathway,
list.gene.snp=list.gene.snp, p.snp.permut=res.permut,
p.snp.obs=res.obs,inspect.snp.n=5,inspect.snp.percent=0.05,
inspect.gene.n=10,inspect.gene.percent=0.05,temp.dir="TEMP/",nperm=90)
result.2

##Third example: Survival data
##observed and permuted p-values have already been computed

data(data.surv)
data(data.pathway.surv)
data(list.gene.snp.surv)
path.data <- paste(system.file("sampleData", package="PIGE"),"/",sep="")
res.permut <- read.table(file=paste(path.data,"res-permut-surv.txt",sep="")
,header=TRUE,sep=" ")
res.obs   <- read.table(file=paste(path.data,"res-obs-surv.txt",sep="")
,header=TRUE,sep=" ")
result.3 <- ARTP.GE(data.gene.pathway=data.pathway.surv,
list.gene.snp=list.gene.snp.surv, p.snp.permut=res.permut,
p.snp.obs=res.obs,inspect.snp.n=5,inspect.snp.percent=0.05,
inspect.gene.n=10,inspect.gene.percent=0.05,temp.dir="TEMP/",nperm=90)
result.3

## End(Not run)
```

`bootstrap.snp` *Parallel computing of the Likelihhod ratio test (or Wald test) for an interaction term (or a simple SNP effect) on bootstrap sample*

Description

The `bootstrap.snp` function performs on bootstrap sample a Likelihood Ratio Test (or a Wald test) for an interaction term SNP*E (where E is an Environment variable) or for the effect of the SNP. This function uses the parallel computing on different CPU of the computer. This function returns a matrix containing for each bootstrap and SNP the p-value of the interaction term tested (SNP*E) or the p-value of the SNP effect tested. Note that this parametric bootsrap approach has been only currently implemented in this package for Logistic model.

Usage

```
bootstrap.snp(model, Outcome.model = "binary", data,
  var.inter = NULL, indice.snp, class.inter = NULL,
  nbcpu = NULL, Nboot = 1000,file.out = "res-boot")
```

Arguments

model	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
Outcome.model	a character string naming the type of outcome considered. The current version allows only "binary" (by default).
data	a data frame containing the variables in the model.
var.inter	name of the variable which is tested in interaction with the SNPs (SNP:E). By default var.inter=NULL correspond to a test on the SNPs (no interaction)
indice.snp	vector or character indicating the SNPs to be tested.
class.inter	class of the var.inter variable. By default, the variable is considered as continuous and a Wald test is performed. Use ("factor") to indicate categorical variable.
nbcpu	integer indicating the number of CPU of your computer (-1). By default, the function use only one cpu.
Nboot	number of bootstrap (1000 by default).
file.out	name of the output file where the result will be saved.

Value

A matrix containing the p-value, for each bootstrap (row) and for each SNP (column), of the likelihood ratio test (or the Wald test) for the interaction term or the SNP effect. This matrix is also saved in a txt file (named by the argument file.out) located in the current directory.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

Examples

```
data(data.pige)
data(data.pathway)
data(list.gene.snp)
res <-data.to.PIGE(data=data.pige,data.pathway=data.pathway,
list.gene.snp=list.gene.snp,choice.pathway=c(1,2))
formul <- formula(y~factor(cov1)+factor(cov2)+factor(cov3)+factor(cov4)
+var_int)
debut <- Sys.time()
p.snp.boot.ex <-  bootstrap.snp(model=formul,data=data.pige,
indice.snp=res$snp.selected,var.inter="var_int",class.inter=NULL,
nbcpu=2,Nboot=2,file.out="res-boot")
print(Sys.time()-debut)
```

`compute.p.snp.obs` *Likelihhod ratio test (or Wald test) for an interaction term (or a single SNP effect)*

Description

The `compute.p.snp.obs` function performs a Likelihood Ratio Test (LRT) for an interaction term SNP*E (where E is an environmental variable) or for the effect of the SNP. The function return the p-value of the Likelihood Ratio Test performed for all the SNPs in the original data set. The SNPs are considered as continuous variable (coded 0,1,2).

Usage

```
compute.p.snp.obs(data, Outcome.model = "binary", model,
  indice.snp, var.inter, class.inter = "NULL",
  file.out = "res-obs")
```

Arguments

`data`	a data frame containing the variables in the model
`model`	an object of class "formula": a symbolic description of the model to be fitted without the interaction term
`Outcome.model`	a character string naming the type of outcome considered. It could be "`binary`" (by default) or "`survival`".
`indice.snp`	vector or character indicating the SNPs to be tested
`var.inter`	name of the variable which is tested in interaction with the SNPs (SNP:E). By default var.inter=NULL correspond to a test on the SNPs (no interaction)
`class.inter`	class of the var.inter variable. By default, the variable is considered continuous and a wald test is performed. Use "`factor`" to indicate a categorical variable.
`file.out`	name of the output file where the result will be saved.

Value

A data frame with one row containing the pvalue of the likelihood ratio test (or Wald test) for the interaction term (or a SNP effect) for all the SNPs. This data frame is also saved in a txt file (named by the argument `file.out`) located in the current directory.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

Examples

```
data(data.pige)
## Case-control study:
data(data.pathway)
data(list.gene.snp)
res <-data.to.PIGE(data=data.pige,data.pathway=data.pathway,
list.gene.snp=list.gene.snp,choice.pathway=c(1,2))
formul <- formula(y~factor(cov1)+factor(cov2)+factor(cov3)+factor(cov4)
+var_int)
p.snp.obs.ex <-  compute.p.snp.obs(data=data.pige,model=formul,
indice.snp=res$snp.selected,var.inter="var_int",class.inter=NULL)

## Survival analysis
data(data.surv)
data(data.pathway.surv)
data(list.gene.snp.surv)
res1 <-data.to.PIGE(data=data.surv,data.pathway=data.pathway.surv,
list.gene.snp=list.gene.snp.surv,choice.pathway=c(1:7))
formul <- formula(Surv(TIME, EVENT) ~ var_int)
p.snp.obs.ex <- compute.p.snp.obs(data=data.surv,Outcome.model="surv"
,model=formul,indice.snp=res1$snp.selected,var.inter="var_int"
,class.inter=NULL,file.out="res-obs-surv")
```

`data.pathway` *Fictive data frame corresponding to the gene x pathway analysis for the example of case-control study.*

Description

Fictive data frame corresponding to the gene x pathway analysis for the example of case-control study.

Format

A data frame of 0 and 1 indicating in which pathway (column) each gene (rows) is included.

`data.pathway.surv` *Fictive data frame corresponding to the gene x pathway analysis for the example of survival analysis.*

Description

Fictive data frame corresponding to the gene x pathway analysis for the example of survival analysis.

Format

A data frame of 0 and 1 indicating in which pathway (column) each gene (rows) is included.

data.pige *Sample data of a case-control study*

Description

This fictive data set contains one environmental variable (var_int), 4 fixed covariates (cov) and SNP variables (coded 0,1,2).

Format

A dataframe with 1000 rows and 134 columns

data.surv *Fictive data set for sample data of a survival analysis.*

Description

This fictive data set contains a time to event (TIME), an indicator variable (named EVENT) coded 0/1 (censored/event), one environmental variable (var_int), and SNP variables.

Format

A dataframe with 66 rows and 887 columns

data.to.list.gene.snp *Annotation gene/SNP*

Description

The `data.to.list.gene.snp` function generates the list containing for each gene the corresponding SNP ids from a txt file containing the SNP/gene annotation data. This list is required for the functions `data.to.PIGE` and `ARTP.GE`.

Usage

```
data.to.list.gene.snp(file, header = TRUE, path = NULL)
```

Arguments

`file`	txt file containing the SNP/gene annotation (see Details for the format of the txt file).
`header`	a logical value indicating whether the file contains the names of the variables as its first line. By default `header` is set to TRUE.
`path`	linking to the directory containing the data (SNP/gene)

Details

The txt file containing the annotation data for the SNP/gene is a two columns matrix. The first colums is the name of each SNP. The second columns indicates the genes which the corresponding SNP (same row) belongs. It will be noted "Gene1/Gene4/Gene5", if for example a SNP belongs to the genes: Gene1, Gene4, Gene5.

Value

A list containing for each gene the names of the SNPs belonging to it. This list is required for using the function data.to.PIGE and the function ARTP.GE.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

Examples

```
##Example : case-control study data
data(data.pige)
data(data.pathway)
path.in <- paste(system.file("sampleData", package="PIGE"),"/",sep="")
file <- "SNP-GENE-annotation.txt"
list.gene.snp <- data.to.list.gene.snp(file,path=path.in)
##Example Survival data
data("data.surv")
data("data.pathway.surv")
path.in <- paste(system.file("sampleData", package="PIGE"),"/",sep="")
file="snp.gene.surv.txt"
list.gene.snp.surv <- data.to.list.gene.snp(file,path=path.in)
```

---

data.to.PIGE *Generate data for Gene- and Pathway-environment interaction*

Description

The data.to.PIGE function prepare your data to be analysed by the functions permutation.snp and compute.p.snp.obs.

Usage

```
data.to.PIGE(data, data.pathway, list.gene.snp,
  choice.pathway = NULL)
```

Arguments

data	a data frame containing the variables in the model and the SNP.
data.pathway	A data frame (Gene X Pathway) of 0 and 1 values. The rownames (gene name considered) and the colnames (names of the studied pathways). have to be specified. The value 1 indicates that a gene is included in the corresponding pathway.
list.gene.snp	List containing for each gene the names of the SNPs belonging to it. This list could be generated by data.to.list.gene.snp function.
choice.pathway	names or indice of the pathway to be analysed.

Value

The returned value is a list containing:

data.pathway	A data frame (Gene X Pathway) of 0 and 1 values similar as data.pathway but only for the selected pathways.
tab.snp.gene	A matrix with in the first column the names of the SNPs and the corresponding gene on the second column. Note that a SNPs could belong to several gene.
snp.selected	A vector of names corresponding to the SNPs analalysed.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

Examples

```
data(data.pathway)
data(data.pige)
data(list.gene.snp)
#Example: one pathway
res1 <-data.to.PIGE(data=data.pige,data.pathway=data.pathway,
list.gene.snp=list.gene.snp,choice.pathway=c(2))
#Example: two pathways
res <-data.to.PIGE(data=data.pige,data.pathway=data.pathway
,list.gene.snp=list.gene.snp,choice.pathway=c(1,2))
```

---

list.gene.snp — *Fictive list for the case-control study example containing the names of the snp for each gene included in the studied pathways*

Description

Fictive list for the case-control study example containing the names of the snp for each gene included in the studied pathways

Format

A list containing the names of the SNPs belonging to each gene analysed.

list.gene.snp.surv — *Fictive list (for survival example) containing the names of the snp for each gene included in the studied pathways*

Description

Fictive list (for survival example) containing the names of the snp for each gene included in the studied pathways

Format

A list containing the names of the SNPs belonging to each gene analysed.

LR.cont — *Wald test for an adjusted model.*

Description

The LR.cont function performs a Wald test for a SNP effect. The function returns the p-value of the Wald test. It is an Internal function used by the permutation.snp function.

Usage

```
LR.cont(x, formula, data)
```

Arguments

x	name or numeric vector corresponding to the SNP tested.
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
data	a data frame containing the variables in the model.

Value

p-value of the Wald test for a SNP effect.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.cont.boot *Wald test for an adjusted model.*

Description

The LR.cont,boot function performs a Wald test for a SNP effect. The function returns the p-value of the Wald test. It is an Internal function used by the bootstrap.snp function.

Usage

```
LR.cont.boot(x, formula, data)
```

Arguments

x	numeric vector corresponding to the new response variable (from parametric bootstrap).
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
data	a data frame containing the variables in the model.

Value

p-value of the Wald test for a SNP effect.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.cont.surv *Wald test for an adjusted model.*

Description

The LR.cont.surv function performs a Wald test for a SNP effect based on a cox model. The function returns the p-value of the Wald test. It is an Internal function used by the permutation.snp function.

Usage

```
LR.cont.surv(x, formula, data)
```

Arguments

x	name or numeric vector corresponding to the SNP tested.
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
data	a data frame containing the variables in the model.

Value

p-value of the Wald test for a SNP effect.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.inter.cat *Likelihhod ratio test for an interaction term*

Description

The LR.inter.cat function performs a likelihood ratio test (LRT) for an interaction term between a categorical variable and a SNP (coded 0,1, 2) in a logistic model. The function returns the p-value of the likelihood ratio test.

Usage

```
LR.inter.cat(x, formula, data, Z1)
```

Arguments

x	name or numeric vector corresponding to the SNP tested.
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
data	a data frame containing the variables in the model.
Z1	name of the variable which is tested in interaction with x (x:Z1).

Value

p-value of the likelihood ratio test for the interaction term.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.inter.cat.boot *Likelihhod ratio test for an interaction term*

Description

The LR.inter.cat.boot function performs a likelihood ratio test (LRT) for an interaction term between a categorical variable and a SNP (coded 0,1, 2) in a logistic model. The function returns the p-value of the likelihood ratio test.

Usage

```
LR.inter.cat.boot(x, formula, data, Z1)
```

Arguments

x	numeric vector corresponding to the new response variable (from parametric bootstrap).
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
data	a data frame containing the variables in the model.
Z1	name of the variable which is tested in interaction with x (x:Z1).

Value

p-value of the likelihood ratio test for the interaction term.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.inter.cat.surv *Likelihhod ratio test for an interaction term*

Description

The LR.inter.cat.surv function performs a likelihood ratio test (LRT) for an interaction term between a categorical variable and a SNP (coded 0,1,2) in a cox model. The function returns the p-value of the likelihood ratio test.

Usage

```
LR.inter.cat.surv(x, formula, data, Z1)
```

Arguments

x	name or numeric vector corresponding to the SNP tested.
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
data	a data frame containing the variables in the model.
Z1	name of the variable which is tested in interaction with x (x:Z1).

Value

p-value of the likelihood ratio test for the interaction term.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.inter.cont *Wald test for an interaction term*

Description

The LR.inter.cont function performs a Wald test for an interaction term between a continuous variable and a SNP. The function returns the p-value of the Wald test.

Usage

```
LR.inter.cont(x, formula, data, Z1)
```

Arguments

x	name or numeric vector corresponding to the SNP tested
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term
data	a data frame containing the variables in the model
Z1	name of the variable which is tested in interaction with x (x:Z1)

Value

pvalue of the Wald test for the interaction term

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.inter.cont.boot *Wald test for an interaction term*

Description

The LR.inter.cont.boot function performs a Wald test for an interaction term between a continuous variable and a SNP. The function returns the p-value of the Wald test.

Usage

```
LR.inter.cont.boot(x, formula, data, Z1)
```

Arguments

x	numeric vector corresponding to the new response variable (from parametric bootstrap)
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term
data	a data frame containing the variables in the model
Z1	name of the variable which is tested in interaction with x (x:Z1)

Value

pvalue of the Wald test for the interaction term

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

LR.inter.cont.surv *Wald test for an interaction term*

Description

The LR.inter.cont.surv function performs a Wald test for an interaction term between a continuous variable and a SNP based on a cox model.. The function return the pvalue of the Wald test.

Usage

```
LR.inter.cont.surv(x, formula, data, Z1)
```

Arguments

x	name or numeric vector corresponding to the SNP tested
formula	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term
data	a data frame containing the variables in the model
Z1	name of the variable which is tested in interaction with x (x:Z1)

Value

pvalue of the Wald test for the interaction term

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

---

permutation.snp *Parallel computing of the Likelihhod ratio test (or Wald test) for an interaction term (or a simple SNP effect) on permutation sample*

Description

The permutation.snp function performs on permutation sample a Likelihood Ratio Test (or a Wald test) for an interaction term SNP*E (where E is an Environment variable) or for the effect of the SNP. This function uses the parallel computing on different CPU of the computer. This function returns a matrix containing for each permutation and SNP the p-value of the interaction term tested (SNP*E) or the p-value of the SNP effect tested.

Usage

```
permutation.snp(model, Outcome.model = "binary", data, var.inter = NULL,
  indice.snp, class.inter = NULL, method = "YX", nbcpu = NULL,
  Npermut = 1000, file.out = "res-permut")
```

Arguments

model	an object of class "formula" : a symbolic description of the model to be fitted without the interaction term.
Outcome.model	a character string naming the type of outcome considered. It could be "binary" (by default) or "survival".
data	a data frame containing the variables in the model.
var.inter	name of the variable which is tested in interaction with the SNPs (SNP:E). By default var.inter=NULL correspond to a test on the SNPs (no interaction)
indice.snp	vector or character indicating the SNPs to be tested.

class.inter	class of the var.inter variable. By default, the variable is considered as continuous and a Wald test is performed. Use ("factor") to indicate categorical variable.
method	method choice for the permutation. By default "YX" a permutation of the phenotype and the adjusted effect are performed otherwise only the phenotype is permuted. A method "YwithinX" permutes the outcome within levels of the "var.inter variable".
nbcpu	integer indicating the number of CPU of your computer (-1). By default, the function use only one cpu.
Npermut	number of permutation (1000 by default).
file.out	name of the output file where the result will be saved.

Value

A matrix containing the p-value, for each permutation (row) and for each SNP (column), of the likelihood ratio test (or the Wald test) for the interaction term or the SNP effect. This matrix is also saved in a txt file (named by the argument file.out) located in the current directory.

Author(s)

Benoit Liquet <benoit.liquet@isped.u-bordeaux2.fr>
Therese Truong <therese.truong@inserm.fr>

Examples

```
data(data.pige)
data(data.pathway)
data(list.gene.snp)
res <-data.to.PIGE(data=data.pige,data.pathway=data.pathway,
list.gene.snp=list.gene.snp,choice.pathway=c(1,2))
formul <- formula(y~factor(cov1)+factor(cov2)+factor(cov3)+factor(cov4)
+var_int)
debut <- Sys.time()
p.snp.permut.ex <-  permutation.snp(model=formul,data=data.pige,
indice.snp=res$snp.selected,var.inter="var_int",class.inter=NULL,nbcpu=2,
Npermut=2,file.out="res-permut")
print(Sys.time()-debut)
## Not run:
##Survival example:
data(data.surv)
data(data.pathway.surv)
data(list.gene.snp.surv)
res1 <-data.to.PIGE(data=data.surv,data.pathway=data.pathway.surv,
list.gene.snp=list.gene.snp.surv,choice.pathway=c(1:7))
formul <- formula(Surv(TIME, EVENT) ~ var_int)
p.snp.permut.ex <-  permutation.snp(model=formul,Outcome.model="surv"
,data=data.surv,indice.snp=res1$snp.selected,var.inter="var_int",
class.inter=NULL,nbcpu=3,Npermut=9,file.out="res-permut-surv")
##YwithinX example:
```

```
data(data.pathway)
data(list.gene.snp)
res <-data.to.PIGE(data=data.pige,data.pathway=data.pathway,
list.gene.snp=list.gene.snp,choice.pathway=c(1,2))
formul <- formula(y~factor(cov1)+factor(cov2)+factor(cov3)+factor(cov4)
+var_int)
debut <- Sys.time()
p.snp.permut.ex <-  permutation.snp(model=formul,data=data.pige, method = "YwithinX",
indice.snp=res$snp.selected,var.inter="var_int",class.inter=NULL,nbcpu=3,
Npermut=9,file.out="res-permut")
print(Sys.time()-debut)

## End(Not run)
```

`snp.gene.select` *Fictive data frame containing the names of the gene (first column) and the names of the SNP belonging to the SNPS (second column)*

Description

Fictive data frame containing the names of the gene (first column) and the names of the SNP belonging to the SNPS (second column)

Format

A data frame of 130 entries corresponding of the gene's name of each SNP.

Index