Package ‘SNPmaxsel’

February 19, 2015

Version 1.0-3

Date 2009-05-29

Title Maximally selected statistics for SNP data

Author Anne-Laure Boulesteix <boulesteix@ibe.med.uni-muenchen.de>

Maintainer Anne-Laure Boulesteix <boulesteix@ibe.med.uni-muenchen.de>

Depends R (>= 2.2.1), mvtnorm, combinat

Suggests mvtnorm, combinat

Description This package implements asymptotic methods related to
maximally selected statistics, with applications to SNP data.

License GPL (>= 2)

URL http://cran.r-project.org/web/packages/SNPmaxsel/index.html

Repository CRAN

Date/Publication 2012-10-29 08:57:40

NeedsCompilation no

R topics documented:

Fasymp .................................................. 2
maxsel .................................................. 3
maxsel.asymp.test .................................. 4
transf.inter ......................................... 6

Index 8
**Description**

This function computes the asymptotical distribution of the maximally selected chi-square statistic, where maximal selection is performed over the test statistics defined by `type` or user-defined through `groups`.

**Usage**

```r
fasymp(t, a.vec, type=NULL, groups=NULL)
```

**Arguments**

- **t**
  - the value at which the distribution function has to be computed.
- **a.vec**
  - A vector of length K giving the proportion of observations in each category. Its sum must be 1. If `type` = "ordinal" or "inter.cat", `a.vec` must have length 9.
- **type**
  - must be one of "ordinal", "all.pairs", "all.partitions", "inter.ord", "inter.cat", "inter.ord.main". If `type`=NULL, the chi-square statistics are computed for the groups defined by `groups`.
- **groups**
  - If `type`=NULL, the chi-square statistics are computed for the groups defined by `groups`. `groups` must be a list with m elements, where m is the number of considered chi-square statistics. Each element is a list with two elements `group1` and `groupR`. `groups$group1` and `groups$groupR` are numeric vectors giving the indices of the categories included in both groups. See example below.

**Details**

This function uses the function `pmvnorm` from the package `mvtnorm`. If the considered test statistics are defined by `groups`, `groups` should not have more than 100 elements, since the `pmvnorm` function becomes unstable (or computationally prohibitive) for such a high dimension.

**Value**

the value of the distribution function at t.

**Author(s)**

Anne-Laure Boulesteix (http://www.ibe.med.uni-muenchen.de/organisation/mitarbeiter/020_professuren/boulesteix/index.html)

**References**

maxsel

See Also
 maxsel.asymp.test, maxsel.

Examples

# load SNPmaxsel library
# library(SNPmaxsel)

Fasymp(t=2,a.vec=c(0.2,0.1,0.05,0.05,0.05),type="ord")
Fasymp(t=2,a.vec=c(0.2,0.1,0.05,0.05,0.05),type="all.pairs")
Fasymp(t=2,a.vec=c(0.2,0.1,0.05,0.05,0.05),type="all.partitions")
Fasymp(t=2,a.vec=c(0.2,0.1,0.1,0.05,0.05,0.05,0.05,0.3),type="inter.ord")
Fasymp(t=2,a.vec=c(0.2,0.1,0.1,0.05,0.05,0.05,0.05,0.3),type="inter.cat")

# Creating a user-defined groups list
# (corresponding to type="ord" with K=5)
my.groups<-list(list(group1=1:groupR=R:U),
list(group1=1:2,groupR=group1:2:5),
list(group1=1:3,groupR=group1:4:5),
list(group1=1:4,groupR=5))
Fasymp(t=2,a.vec=c(0.2,0.1,0.05,0.05,0.05,0.3),type=PhysNull,groups=my.groups)

maxsel Computes maximally selected chi-square statistics

Description

The function maxsel computes the maximal chi-square statistic over some candidate binary splits specified by type.

Usage

maxsel(x1,x2=NULL,y,type="inter.ord")

Arguments

x1 a numeric vector of length n giving the values of the variable x1 for the considered n observations. The classes must be coded as 1,...,K.

x2 a numeric vector of length n giving the values of the variable x2 for the considered n observations. x2 should be NULL (default) for all types other than "inter.ord", "inter.cat" and "inter.ord.main". Since interactions are implemented for K=3 only (SNPs), x2 must be coded as 1,2,3.

ty a numeric vector of length n giving the class (response variable Y) of the considered n observations. The classes must be coded as 0 and 1.

type must be one of "ordinal", "all.pairs", "all.partitions", "inter.ord", "inter.cat", "inter.ord.main".

Value

the value of the maximally selected chi-square statistic.

Author(s)

Anne-Laure Boulesteix (http://www.ibe.med.uni-muenchen.de/organisation/mitarbeiter/020_professuren/boulesteix/index.html)

References


See Also

maxsel.asymp.test, Fasymp.

Examples

# load SNPmaxsel library
# library(SNPmaxsel)

x1<-sample(5,1000,replace=TRUE)
y<-sample(c(0,1),1000,replace=TRUE)
maxsel(x1=x1,y=y,type="ord")
maxsel(x1=x1,y=y,type="all.pairs")
maxsel(x1=x1,y=y,type="all.partitions")

x1<-sample(3,1000,replace=TRUE)
x2<-sample(3,1000,replace=TRUE)
maxsel(x1=x1,x2=x2,y=y,type="inter.ord")
maxsel(x1=x1,x2=x2,y=y,type="inter.cat")
maxsel(x1=x1,x2=x2,y=y,type="inter.ord.main")

maxsel.asymp.test

Test of independence based on maximally selected statistics

Description

The function maxsel.test computes the probability that the maximally selected chi-square statistic is <= than the value observed from the data, under the null-hypothesis of no association between X and Y, given the proportions of observations with X=1,...,X=K. The candidate binary splits over which the chi-square statistic is maximized is specified by type. If x denotes the output of the function maxsel.asymp.test, 1-x may be seen as the p-value of an independence test.
maxsel.asymp.test

Usage

maxsel.asymp.test(x1, x2=NULL, y, type)

Arguments

x1 a numeric vector of length n giving the values of the variable x1 for the considered n observations. The classes must be coded as 1,...,K.

x2 a numeric vector of length n giving the values of the variable x2 for the considered n observations. x2 should be NULL (default) for all types other than "inter.ord", "inter.cat" and "inter.ord.main". Since interactions are implemented for K=3 only (SNPs), x2 must be coded as 1,2,3.

y a numeric vector of length n giving the class (response variable Y) of the considered n observations. The classes must be coded as 0 and 1.

type must be one of "ordinal", "all.pairs", "all.partitions", "inter.ord", "inter.cat", "inter.ord.main".

Details


Value

maxselstat the observed maximally selected statistic.

value the value of the distribution function of the maximally selected statistic at maxselstat.

Author(s)

Anne-Laure Boulesteix (http://www.ibe.med.uni-muenchen.de/organisation/mitarbeiter/020_professuren/boulesteix/index.html)

References


See Also

maxsel, fasymp.

Examples

# load SNPmaxsel library
# library(SNPmaxsel)

x1<-sample(5,1000,replace=TRUE)
y<-sample(c(0,1),1000,replace=TRUE)

maxsel.asymp.test(x1=x1,y=y,type="ord")
maxsel.asymp.test(x1=x1,y=y,type="all.pairs"
maxsel.asymph.test(x1=x1,y=y,type="all.partitions")

x1<-sample(3,1000,replace=TRUE)
x2<-sample(3,1000,replace=TRUE)

maxsel.asymph.test(x1=x1,x2=x2,y=y,type="inter.ord")
maxsel.asymph.test(x1=x1,x2=x2,y=y,type="inter.cat")
maxsel.asymph.test(x1=x1,x2=x2,y=y,type="inter.ord.main")

---

transf.inter

Transforms a pair of SNPs into a single variable with nine categories

Description

The function transf.inter creates a new variable with 9 categories out of a pair of variables with 3 categories (e.g. SNPs). The new variable is generated as described in Boulesteix et al (2007): =1 if x1=1 and x2=1, =2 if x1=2 and x2=1, =3 if x1=3 and x2=1, =4 if x1=1 and x2=2, =5 if x1=2 and x2=2, =6 if x1=3 and x2=2, =7 if x1=1 and x2=3, =8 if x1=2 and x2=3, =9 if x1=3 and x2=3.

Usage

transf.inter(x1,x2)

Arguments

x1 a numeric vector of length n giving the first SNP, coded as 1,2,3.
x2 a numeric vector of length n giving the second SNP, coded as 1,2,3.

Value

a numeric vector of length n containing the new variable with 9 categories.

Author(s)

Anne-Laure Boulesteix (http://www.ibe.med.uni-muenchen.de/organisation/mitarbeiter/020_professuren/boulesteix/index.html)

References


See Also

cmaxsel.maxsel.asymph.test.
Examples

# load SNPmaxsel library
# library(SNPmaxsel)

x1<-sample(3,100,replace=TRUE)
x2<-sample(3,100,replace=TRUE)

transf.inter(x1,x2)
Index

*Topic htest
  Fasymp, 2
  maxsel, 3
  maxsel.asymp.test, 4
  transf.inter, 6

Fasymp, 2, 4, 5
maxsel, 3, 3, 5, 6
maxsel.asymp.test, 3, 4, 4, 6
transf.inter, 6