

Package ‘pleio’

December 7, 2018

Type Package

Title Pleiotropy Test for Multiple Traits on a Genetic Marker

Version 1.6

Date 2018-12-06

Author Daniel J Schaid

Maintainer Jason Sinnwell <sinnwell.jason@mayo.edu>

Description Perform tests for pleiotropy of multiple traits of various variable types on genotypes for a genetic marker.

License GPL (>= 2)

Depends R (>= 3.2.0), rms, Matrix, stats, utils

Suggests R.rsp

VignetteBuilder R.rsp

URL <http://bioinformaticstools.mayo.edu/research/pleio/>

NeedsCompilation no

Repository CRAN

Date/Publication 2018-12-06 23:50:03 UTC

R topics documented:

pleio.demo	2
pleio.glm.fit	2
pleio.glm.sequential	4
pleio.glm.test	5
pleio.q.fit	6
pleio.q.sequential	7
pleio.q.test	9
pleio.qdemo	10

Index	11
--------------	-----------

pleio.demo	<i>Demonstration dataset for pleiotropy tests with a mixture of trait types and covariates</i>
------------	--

Description

A simulated dataset with 3 y variables simulated from multivariate normal with common correlation structure, correlation of 0.5. The traits are (1) gaussian, (2) binomial, and (3) ordinal. Genotypes were simulated based on minor allele frequency of 0.2, and assume that the 3 traits are not associated with dose of minor allele (all 3 betas = 0).

Usage

```
data("pleio.demo")
```

Format

y is a data.frame with 500 observations and 3 traits, whose distribution families are gaussian, binomial, and ordinal, respectively; x is a matrix of 5 covariates, geno is a single SNP of dosage of a minor allele simulated from frequency of 0.2.

Examples

```
data(pleio.demo)
str(y)
table(geno)
x[1:5,]
```

pleio.glm.fit	<i>Compute Intermediate Statistics for Testing Pleiotropy</i>
---------------	---

Description

Pre-compute statistics for testing pleiotropy, to be used in pleio.glm.test and pleio.glm.sequential

Usage

```
pleio.glm.fit(y, g, glm.family, x.all=NULL, x.index.list=NULL)
```

Arguments

<code>y</code>	matrix of multiple traits to test for pleiotropy on genotype vector. Rows are for subjects and columns are for traits.
<code>g</code>	vector of genotypes coded as dosage of a minor allele, 0,1, and 2. Allows imputed genotype dosage.
<code>glm.family</code>	vector of strings corresponding to the type of phenotype for the columns of <code>y</code> . Types accepted are gaussian, binomial, and ordinal.
<code>x.all</code>	matrix of all covariates traits to adjust for the traits.
<code>x.index.list</code>	a list of vectors containing the index of the columns of <code>x.all</code> for which to be adjusted for each trait. The list must have the same number of vectors as there are traits in <code>y</code> . If one trait has no covariates, specify a vector containing 0 for that trait.

Details

The matrix `y` of traits is expected to have rows for subjects and columns for traits. The number of rows of `y` should match the length of vector `g`. If adjusting traits by covariates in `x.all`, the `x.index.list` must be a list containing a vector of indices of the columns of `x.all` for each trait. If one trait does not have covariates and others do, then the vector for the trait with no covariate must be a vector containing the number 0. This function computes intermediate statistics that are subsequently used to test various hypotheses about pleiotropy. For technical details, Schaid et al. (2016, 2017)

Value

A list containing	
<code>theta</code>	
<code>n.intercepts</code>	
<code>n.coef.covar</code>	
<code>n.parm</code>	
<code>an.mat</code>	
<code>n.traits</code>	number of traits (columns) in <code>y</code>

Author(s)

Dan Schaid, Jason Sinnwell.

References

- Schaid DJ, Tong X, Larrabee B, Kennedy RB, Poland GA, Sinnwell JP. Statistical Methods for Testing Genetic Pleiotropy. *Genetics*. 2016 Oct;204(2):483-497.
- Schaid DJ, Tong X, Batzler A, Sinnwell JP, Qing J, Biernacka JM. Multivariate Generalized Linear Model for Genetic Pleiotropy. Under review.

pleio.glm.sequential *Perform sequential tests of pleiotropy*

Description

Perform sequential tests of pleiotropy to determine the number of traits, and which traits, are associated with a vector of genotypes.

Usage

```
pleio.glm.sequential(obj.pleio.glm.fit, pval.threshold)
```

Arguments

`obj.pleio.glm.fit`
 result of `pleio.glm.fit`

`pval.threshold` p-value for rejecting the null hypothesis of the specified number of coefficients constrained to be zero.

Details

Perform sequential tests of pleiotropy, starting at the usual multivariate null hypothesis that all coefficients = 0. If this test rejects because the p-value < `pval.threshold`, then allow one coefficient to be non-zero in order to test whether the remaining coefficients = 0. If the test of one non-zero coefficient rejects, then allow two non-zero coefficients, considering all possible combinations of two non-zero coefficients and test whether the remaining coefficients = 0. Continue this sequential testing until the p-value for a test is greater than the specific `pval.threshold`. The step at which the p-value > `pval.threshold` determines which traits are associated with the genotype. If there are `m` traits, the sequential testing stops either when p-value > `pval.threshold`, or when (`m-1`) traits are tested. If the p-value remains less than `pval.threshold` when testing (`m-1`) traits, this implies that all `m` traits are associated with the genotype.

Value

A list containing:

`pval` p-value of the final test from the sequential testing

`count` the number of nonzero coefficients

`index.nonzero.coef`
 index of column(s) of `y` that have non-zero coefficients. These indices indicate which traits are associated with the genotype, accounting for the correlations among the traits.

Author(s)

Dan Schaid and Jason Sinnwell

References

Schaid DJ, Tong X, Larrabee B, Kennedy RB, Poland GA, Sinnwell JP. Statistical Methods for Testing Genetic Pleiotropy. *Genetics*. 2016 Oct;204(2):483-497.

Schaid DJ, Tong X, Batzler A, Sinnwell JP, Qing J, Biernacka JM. Multivariate Generalized Linear Model for Genetic Pleiotropy. Under review.

Examples

```
data(pleio.demo)

## test without covars
fams <- c("gaussian", "binomial", "ordinal")
obj <- pleio.glm.fit(y, geno, glm.family=fams)
stat <- pleio.glm.test(obj, count.nonzero.coef = 0)
stat$stat
stat$pval
pseq <- pleio.glm.sequential(obj, pval.threshold=.5)
pseq
```

pleio.glm.test

Single test of the number of traits associated with genotype

Description

Perform single test of the number of traits associated with a genotype, by allowing a specified number of regression beta's to be unconstrained.

Usage

```
pleio.glm.test(obj.pleio.glm.fit, count.nonzero.coef = 0)
```

Arguments

`obj.pleio.glm.fit`

result of `pleio.glm.fit`

`count.nonzero.coef`

Number of non-zero coefficients (betas). A non-zero beta is allowed to be unconstrained, while all other beta's are constrained to be zero.

Details

By specifying the number of non-zero coefficients, the algorithm evaluates all possible ways of choosing unconstrained and constrained betas, and for each configuration a statistic (tk) is computed. This tk statistic can be considered a measure of fit of a model. The minimum tk over all possible configurations provides a global test of whether one of the models fits well.

Value

A list containing:

stat	global test statistic
df	degrees of freedom of the statistic
pval	p-value for the test
index.nonzero.coef	index of the non-zero coefficients that provide the minimum tk goodness of fit statistic - this configuration is assumed to have coefficients for all other indices equal to zero.
tk	test testistic for the tests performed on trait combinations in vk.set
vk.set	data.frame containing the tests performed. For m traits, and k = count.nonzero.coef, there are m-choose-k tests considered in the null hypothesis. The data.frame provides the indices of the unconstrained coefficients for the corresponding tk test statistic for the configuration. Rows are the indices for each configuration, and the columns are for the different configurations tested.

Author(s)

Dan Schaid and Jason Sinnwell

References

Schaid DJ, Tong X, Larrabee B, Kennedy RB, Poland GA, Sinnwell JP. Statistical Methods for Testing Genetic Pleiotropy. *Genetics*. 2016 Oct;204(2):483-497.

Examples

```
data(pleio.demo)
obj <- pleio.glm.fit(y, geno, glm.family=c("gaussian","binomial","ordinal"))
test1 <- pleio.glm.test(obj, count.nonzero.coef = 0)
test1
test2 <- pleio.glm.test(obj, count.nonzero.coef = 1)
test2
```

pleio.q.fit

Compute Intermediate Statistics for Testing Pleiotropy

Description

Pre-compute statistics for testing pleiotropy, to be used in pleio.q.test and pleio.q.sequential. This is an earlier function for gaussian traits without covariates. The function pleio.glm.fit is more general, allowing for gaussian traits, and possibly adjusting covariates.

Usage

```
pleio.q.fit(y, geno)
```

Arguments

y	matrix of multiple quantitative traits to test for pleiotropy on genotype vector. Rows are for subjects and columns are for traits.
geno	vector of genotypes coded as dosage of a minor allele, 0,1,2. Allows imputed genotype dosage.

Details

The matrix y of traits is expected to have rows for subjects and columns for traits. The number of rows of y should match the length of vector geno. This function computes intermediate statistics that are subsequently used to test various hypotheses about pleiotropy. For technical details, Schaid et al. (Genetics, 2016)

Value

A list containing	
x	the de-correlated genotype design matrix
xx.inv	inverse of $x'x$
beta.ols	estimated unconstrained beta coefficients for y traits based on ordinary least squares
n.traits	number of traits (columns) in y

Author(s)

Dan Schaid, Jason Sinnwell.

References

Schaid DJ, Tong X, Larrabee B, Kennedy RB, Poland GA, Sinnwell JP. Statistical Methods for Testing Genetic Pleiotropy. Genetics. 2016 Oct;204(2):483-497.

pleio.q.sequential *Perform sequential tests of pleiotropy*

Description

Perform sequential tests of pleiotropy to determine the number of traits, and which traits, are associated with a vector of genotypes. For gaussian traits without covariates.

Usage

```
pleio.q.sequential(obj.fit, pval.threshold)
```

Arguments

`obj.fit` result of `pleio.q.fit`

`pval.threshold` p-value for rejecting the null hypothesis of the specified number of non-zero betas.

Details

Perform sequential tests of pleiotropy, starting at the usual multivariate null hypothesis that all coefficients = 0. If this test rejects because the p-value < `pval.threshold`, then allow one coefficient to be non-zero in order to test whether the remaining coefficients = 0. If the test of one non-zero coefficient rejects, then allow two non-zero coefficients, considering all possible combinations of two non-zero coefficients and test whether the remaining coefficients = 0. Continue this sequential testing until the p-value for a test is greater than the specific `pval.threshold`. The step at which the p-value > `pval.threshold` determines which traits are associated with the genotype. If there are `m` traits, the sequential testing stops either when p-value > `pval.threshold`, or when (`m-1`) traits are tested. If the p-value remains less than `pval.threshold` when testing (`m-1`) traits, this implies that all `m` traits are associated with the genotype.

Value

A list containing:

`pval` p-value of the final test from the sequential testing

`index.beta` index of columns of `y` that have non-zero betas. These indices indicate which traits are associated with the genotype, accounting for the correlations among the traits.

Author(s)

Dan Schaid and Jason Sinnwell

References

Schaid DJ, Tong X, Larrabee B, Kennedy RB, Poland GA, Sinnwell JP. Statistical Methods for Testing Genetic Pleiotropy. To appear in *Genetics*, 2016.

Examples

```
data(pleio.qdemo)
fit <- pleio.q.fit(y, geno)
test.seq <- pleio.q.sequential(fit, pval.threshold=.05)
test.seq
```

pleio.q.test

Single test of the number of traits associated with genotype

Description

Perform single test of the number of traits associated with a genotype, by allowing a specified number of regression beta's to be unconstrained.

Usage

```
pleio.q.test(obj.fit, count.nonzero.beta = 0)
```

Arguments

`obj.fit` result of pleio.q.fit
`count.nonzero.beta` Number of non-zero betas. A non-zero beta is allowed to be unconstrained, while all other beta's are constrained to be zero.

Details

By specifying the number of non-zero beta's, the algorithm evaluates all possible ways of choosing unconstrained and constrained beta's, and for each configuration a statistic (tk) is computed. This tk statistic can be considered a measure of fit of a model. The minimum tk over all possible configurations provides a global test of whether one of the models fits well.

Value

A list containing:

`stat` global test statistic
`df` degrees of freedom of the statistic
`pval` p-value for the test
`index.nonzero.beta` index of the non-zero beta(s) that provide(s) the minimum tk goodness of fit statistic - this configuration is assumed to have beta's for all other indices equal to zero.
`tests` data.frame containing the tests performed. For m traits, and k = count.nonzero.beta, there are m-choose-k tests considered in the null hypothesis. The data.frame provides the indices of the unconstrained betas and the corresponding tk test statistic for the configuration.

Author(s)

Dan Schaid and Jason Sinnwell

References

Schaid DJ, Tong X, Larrabee B, Kennedy RB, Poland GA, Sinnwell JP. Statistical Methods for Testing Genetic Pleiotropy. *Genetics*. 2016 Oct;204(2):483-497.

Examples

```
data(pleio.qdemo)
fit <- pleio.q.fit(y, geno)
## usual multivariate test of whether all betas = 0
test0 <- pleio.q.test(fit, count.nonzero.beta = 0)
test0
## test whether allowing 2 betas to be non-zero fits data
test2 <- pleio.q.test(fit, count.nonzero.beta = 2)
test2
```

pleio.qdemo

Demonstration dataset for quantitative pleiotropy tests

Description

A simulated dataset with 6 y variables simulated from multivariate normal distribution with common correlation structure, correlation of 0.5, and genotypes simulated based on minor allele frequency of 0.2, and assume that betas for traits 2 and 3 have non-zero values, while all other traits are not associated with dose of minor allele.

Usage

```
data("pleio.qdemo")
```

Format

y is a data.frame with 500 observations (rows) and 6 quantitative traits (columns), while geno is a single SNP of dosage of a minor allele, simulated from frequency of 0.2.

Examples

```
data(pleio.qdemo)
str(y)
table(geno)
```

Index

*Topic **datasets**

pleio.demo, [2](#)

pleio.qdemo, [10](#)

geno (pleio.demo), [2](#)

pleio.demo, [2](#)

pleio.glm.fit, [2](#)

pleio.glm.sequential, [4](#)

pleio.glm.test, [5](#)

pleio.q.fit, [6](#)

pleio.q.sequential, [7](#)

pleio.q.test, [9](#)

pleio.qdemo, [10](#)

x (pleio.demo), [2](#)

y (pleio.demo), [2](#)