Package 'GAMP'

May 10, 2013

Type Package

Title Global Analysis of Methylation Profiles

Version 0.11

Date 2013-03-17

Author Michael C. Wu

Maintainer Michael Wu <mwu@bios.unc.edu>

Description This package contains the TestCDFs and TestDensities functions for testing whether an outcome variable is associated with the overall profile/distribution of methylation values by approximating the CDF or Density of the methylation values for each individual and then using a variance component test to assess significance. This package is still under development and subject to change.

License GPL(>=2)

Depends R (>= 2.15.0), fda, SKAT (>= 0.82)

R topics documented:

TestCDFs   *Test CDFs*

Description

Tests whether the CDFs of the observed methylation distributions for each individual are associated with an outcome variable.

Usage

```
TestCDFs(Z, y, X = NULL, outcomeType = "C", histBreaks = 1000, numBases = min(histBreaks/4, 35),
```

Arguments

Z	Matrix of methylation values for the individuals in the study. Each row is a different CpG and each column corresponds to a different individual. Missing values are allowed, but omitted for each individual.
y	Vector of outcome variables for each individual in the study. Each element of the vector should be numeric. Note that for dichotomous outcomes y should be 0 or 1.
X	Matrix of possible additional covariates for which adjustment is necessary. Set to NULL of none.
outcomeType	"C" for continuous or "D" for dichotomous outcomes (y).
histBreaks	Number of points to estimate the ECDF. Assumed to be evenly spaced in [0,1]. Defaults to 1000.
numBases	Number of knots in computing the basis. Assumed to be evenly in [0,1]. Defaults to min(histBreaks/4, 35).
lambdas	Grid of tuning parameters for the B-Spline smoothing over which we choose to search. 0 correspond to no smoothing while a larger value tends to smooth more.
kernel	Kernel to be used in the testing stage. In general, the linear kernel is reasonable, but alternative kernels are possible. See SKAT package manual for details.
hideProgress	Boolean describing whether the progress of the function should be output. Defaults to FALSE.
adjustmentDichot	
	For dichotomous outcomes, the variance component test can **sometimes** be conservative. In general, this is not a problem for our setting, but set this to TRUE if adjustments should be made for this possible conservatism.
knotLocs	Points between 0 and 1 at which the knots should be placed for B-spline estimation.

Value

p-value for association.

Note

This function is still under development and subject to change.

Author(s)

Michael C. Wu

See Also

TestDensities

Examples

```
# Simulate data set under null
Z1 = matrix(rbeta(10000*20, 0.48, 0.50), ncol = 20) # simulate 20 cases
Z2 = matrix(rbeta(10000*20, 0.48, 0.50), ncol = 20) # simulate 20 controls
Z = cbind(Z1, Z2)
```

```
y = rep(c(1,0), each = 20) # simulate outcome
TestCDFs(Z,y,X = NULL)

# Simulate data set under alternative
Z1 = matrix(rbeta(10000*20, 0.48, 0.50), ncol = 20) # simulate 20 cases
Z2 = matrix(rbeta(10000*20, 0.50, 0.48), ncol = 20) # simulate 20 controls
Z = cbind(Z1, Z2)
y = rep(c(1,0), each = 20) # simulate outcome
TestCDFs(Z,y,X = NULL)
```

TestDensities   *Test Densities*

Description

Tests whether the densities of the observed methylation distributions for each individual are associated with an outcome variable.

Usage

```
TestDensities(Z, y, X = NULL, outcomeType = "C", histBreaks = 200, lambdas = c(0, exp(-10:10)),
```

Arguments

Z	Matrix of methylation values for the individuals in the study. Each row is a different CpG and each column corresponds to a different individual. Missing values are allowed, but omitted for each individual.
y	Vector of outcome variables for each individual in the study. Each element of the vector should be numeric. Note that for dichotomous outcomes y should be 0 or 1.
X	Matrix of possible additional covariates for which adjustment is necessary. Set to NULL of none.
outcomeType	"C" for continuous or "D" for dichotomous outcomes (y).
histBreaks	Number of breaks for estimating the relative histogram. Assumed to be evenly spaced in [0,1]. Defaults to 200.
lambdas	Grid of tuning parameters for the B-Spline smoothing over which we choose to search. 0 correspond to no smoothing while a larger value tends to smooth more.
kernel	Kernel to be used in the testing stage. In general, the linear kernel is reasonable, but alternative kernels are possible. See SKAT package manual for details.
hideProgress	Boolean describing whether the progress of the function should be output. Defaults to FALSE.
adjustmentDichot	
	For dichotomous outcomes, the variance component test can **sometimes** be conservative. In general, this is not a problem for our setting, but set this to TRUE if adjustments should be made for this possible conservatism.
knotLocs	Locations between 0 and 1 at which the B-spline knots should be placed.

Value

p-value for association.

Note

This function is still under development and subject to change.

Author(s)

Michael C. Wu

See Also

[TestCDFs](#)

Examples

```
# Simulate data set under null
Z1 = matrix(rbeta(10000*20, 0.48, 0.50), ncol = 20) # simulate 20 cases
Z2 = matrix(rbeta(10000*20, 0.48, 0.50), ncol = 20) # simulate 20 controls
Z = cbind(Z1, Z2)
y = rep(c(1,0), each = 20) # simulate outcome
TestDensities(Z,y,X = NULL)

# Simulate data set under alternative
Z1 = matrix(rbeta(10000*20, 0.48, 0.50), ncol = 20) # simulate 20 cases
Z2 = matrix(rbeta(10000*20, 0.50, 0.48), ncol = 20) # simulate 20 controls
Z = cbind(Z1, Z2)
y = rep(c(1,0), each = 20) # simulate outcome
TestDensities(Z,y,X = NULL)
```

Index