Package 'MetaHD'

June 7, 2024

Type Package

Title A Multivariate Meta-Analysis Model for Metabolomics Data

Version 0.1.0

Maintainer Jayamini Liyanage <j.liyanage@latrobe.edu.au>

Author Jayamini Liyanage [aut, cre],
Alysha De Livera [aut]

Description Performs multivariate meta-analysis for high-dimensional metabolomics data for integrating and collectively analysing individual-level metabolomics data generated from multiple studies as well as for combining summary estimates. This approach accounts for correlation between metabolites, considers variability within and between studies, handles missing values and uses shrinkage estimation to allow for high dimensionality.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.3.1

Imports corpcor, Matrix, matrixcalc, Rcpp (>= 1.0.0)

LinkingTo Rcpp, RcppArmadillo

NeedsCompilation yes

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2024-06-07 14:20:09 UTC

Contents

MetaHD *A Multivariate Meta-Analysis Model for Metabolomics Data*

Description

The MetaHD function performs a multivariate meta-analysis for combining summary estimates obtained from multiple metabolomic studies by using restricted maximum likelihood estimation. Assuming a meta-analysis is based on N outcomes/metabolites and K studies:

Arguments

`Y`	: treatment effect sizes of the outcomes. This should be in the form of a K x N matrix
`Slist`	: K-dimensional list of N x N matrices representing within-study variances and covariances of the treatment effects
`Psi`	: N x N matrix representing between-study variances and covariances of the treatment effects. (optional, if not specified this will be estimated internally by "MetaHD" using "estimateBSvar" and "estimateCorMat" functions in "MetaHD" package
`shrinkCor`	: a logical value indicating whether a shrinkage estimator should be used to estimate between-study correlation matrix. Default is TRUE
`method`	: estimation method: "fixed" for fixed-effects models,"reml" for random-effects models fitted through restricted maximum likelihood
`bscov`	: a character vector defining the structure of the random-effects covariance matrix. Among available covariance structures, the user can select "unstructured" to obtain between-study covariance matrix with diagonal elements (variances) estimated using restricted maximul likelihood and off-diagonal elements (covariances) reflecting the correlations estimated via shrinkage and "diag" (diagonal) for between-study variances as diagonal elements and zero co-variances
`rigls.maxiter`	: maximum number of iterations of the restricted iterative generalized least square algorithm. Default is set to 1
`impute.na`	: a logical value indicating whether missing values need to be imputed or not. Default is FALSE
`impute.var`	: multiplier for replacing the missing variances in Slist.(a large value, default is 10^4)

Value

A list of objects containing estimate : a N-dimensional vector of the combined estimates, std.err : a N-dimensional vector of the associated standard errors, pVal : a N-dimensional vector of the p-values, I2.stat : I2 statistic

realdata *An Individual-Level Metabolomics Dataset*

Description

This dataset consists of a list of three data frames containing individual-level data, treatment effect estimates, and their associated variances.

Usage

```
realdata
```

Format

A list of data frames as follows:

`all` A dataframe with 12 rows and 14 columns containing individual-level data of 14 metabolites in columns and a total of 12 samples from two groups that have been run separately in two different instruments, leading to two separate studies in rows.

`effects` A dataframe with 2 rows and 14 columns, containing treatment effect estimates of the 14 metabolites for the 2 studies.

`var` A dataframe with 2 rows and 14 columns, containing associated variances of the treatment effects.

Examples

```
head(realdata$all)
head(realdata$effects)
head(realdata$var)
```

simdata.1 *Simulated Dataset 1 : With Complete Data*

Description

This dataset consists of a list of three data frames containing treatment effect estimates, within-study variances and upper trainangular elements of the within-study correlation matrices.

Usage

```
simdata.1
```

Format

A list of data frames as follows:

`effects` A dataframe with 12 rows and 30 columns, containing simulated treatment effect estimates of 30 metabolites for 12 studies.

`wsvar` A dataframe with 12 rows and 30 columns, containing simulated within-study variances of 30 metabolites for 12 studies.

`wscor` A dataframe with 12 rows and 435 columns, containing the upper traingular elements of simulated within-study correlation matrices for 12 studies.

Examples

```
head(simdata.1$effects)
head(simdata.1$wsvar)
head(simdata.1$wscor)
```

`simdata.2` *Simulated Dataset 2 : With Data Missing-At-Random*

Description

This dataset consists of a list of three data frames containing treatment effect estimates and within-study variances with missing values and upper trainangular elements of the within-study correlation matrices.

Usage

```
simdata.2
```

Format

A list of data frames as follows:

`effects` A dataframe with 12 rows and 30 columns, containing simulated treatment effect estimates of 30 metabolites for 12 studies with missing values.

`wsvar` A dataframe with 12 rows and 30 columns, containing simulated within-study variances of 30 metabolites for 12 studies with missing values.

`wscor` A dataframe with 12 rows and 435 columns, containing the upper traingular elements of simulated within-study correlation matrices for 12 studies.

Examples

```
head(simdata.2$effects)
head(simdata.2$wsvar)
head(simdata.2$wscor)
```

Index