Package 'Elja'

July 3, 2023

Type Package

Title Linear, Logistic and Generalized Linear Models Regressions for the EnvWAS/EWAS Approach

Version 1.0.0

Maintainer Marwan El Homsi <marwan.el-homsi@inserm.fr>

Description Tool for Environment-Wide Association Studies (EnvWAS / EWAS) which are repeated analysis. It includes three functions. One function for linear regression, a second for logistic regression and a last one for generalized linear models.

Depends R (>= 4.3)

Imports stats, devtools, dplyr, ggplot2, MASS

Suggests knitr, rmarkdown, mlbench

License GPL (>= 3)

URL https://github.com/EHMarwan/Elja

BugReports https://github.com/EHMarwan/Elja/issues

VignetteBuilder knitr

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation no

Author Marwan El Homsi [aut, cre, cph]
(<https://orcid.org/0000-0002-5782-6757>, Desbrest Institute of Epidemiology and Public Health, Univ Montpellier, INSERM, Montpellier, France.),
Isabella Annesi-Maesano [ctb, ths, fnd]
(<https://orcid.org/0000-0002-6340-9300>, Desbrest Institute of Epidemiology and Public Health, Univ Montpellier, INSERM, Montpellier, France., Department of Allergic and Respiratory Diseases, Montpellier University Hospital, Montpellier.)

Repository CRAN

Date/Publication 2023-07-03 16:20:05 UTC

R topics documented:

`ELJAglm` *Generalized Linear Models regression for EnvWAS/EWAS analysis*

Description

A tool for Environment-Wide Association Studies (EnvWAS / EWAS) which are repeated analysis. This function is espacially for generalized linear models 'glm' and allows the addition of adjustment variables.

Usage

```
ELJAglm(
  var,
  var_adjust = NULL,
  family = binomial(link = "logit"),
  data,
  manplot = TRUE,
  nbvalmanplot = 100,
  Bonferroni = FALSE,
  FDR = FALSE,
  manplotsign = FALSE
)
```

Arguments

`var`	A categorical and binary variable. It is generally your outcome.
`var_adjust`	A vector containing the names of the fixed adjustment variables for all the models.
`family`	The family and the link use for the glm function.
`data`	A dataframe containing all the variables needed for the analysis.
`manplot`	Generate a Manhattan plot of the results of the analysis.
`nbvalmanplot`	The number of variables to include in each Manhattan plot.
`Bonferroni`	Add a dashed bar to the Manhattan plot showing the Bonferroni significance threshold.
`FDR`	Add a dashed bar to the Manhattan plot showing the False Discovery Rate (Benjamini-Hochberg method) significance threshold. NA if all p-values > FDR corrected p-values.
`manplotsign`	Generates a Manhattan plot with only significant results (p<0.05).

Value

A Dataframe with results for each variable of the model.

References

Dunn OJ. Multiple Comparisons Among Means. Journal of the American Statistical Association. 1961;56(293):52-64. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289-300. MLBench · Distributed Machine Learning Benchmark. Available from: https://mlbench.github.io/ Smith JW, Everhart JE, Dickson WC, Knowler WC, Johannes RS. Using the ADAP Learning Algorithm to Forecast the Onset of Diabetes Mellitus. Proc Annu Symp Comput Appl Med Care. 1988 Nov 9;261–5.

Examples

```
### Loading the PIMA dataset contained in the mlbench package

library(mlbench)
data(PimaIndiansDiabetes)

### Using ELJAlinear to perform EWAS analysis

ELJAglm(var = 'diabetes',data = PimaIndiansDiabetes,
family = binomial(link = "logit"), manplot = TRUE, Bonferroni = TRUE,
FDR = TRUE, nbvalmanplot = 30, manplotsign = FALSE)
results
```

`ELJAlinear` *Linear regression for EnvWAS/EWAS analysis*

Description

A tool for Environment-Wide Association Studies (EnvWAS / EWAS) namely repeated analyses allowing to estimate the relationships between several environmental factors and a health events. This function is especially for linear regressions and allows the addition of adjustment variables.

Usage

```
ELJAlinear(
  var,
  var_adjust = NULL,
  data,
  manplot = TRUE,
  nbvalmanplot = 100,
  Bonferroni = FALSE,
  FDR = FALSE,
  manplotsign = FALSE
)
```

Arguments

var	A categorical and binary variable. It is generally your outcome.
var_adjust	A vector containing the names of the fixed adjustment variables for all the models.
data	A dataframe containing all the variables needed for the analysis.
manplot	Generate a Manhattan plot of the results of the analysis.
nbvalmanplot	The number of variables to include in each Manhattan plot.
Bonferroni	Add a dashed bar to the Manhattan plot showing the Bonferroni significance level.
FDR	Add a dashed bar to the Manhattan plot showing the False Discovery Rate (Benjamini-Hochberg method) significance threshold. NA if all p-values > FDR corrected p-values.
manplotsign	Generates a Manhattan plot with only significant results (p<0.05).

Value

A Dataframe with results for each variable of the model.

References

Dunn OJ. Multiple Comparisons Among Means. Journal of the American Statistical Association. 1961;56(293):52-64. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289-300. MLBench · Distributed Machine Learning Benchmark. Available from: https://mlbench.github.io/ Smith JW, Everhart JE, Dickson WC, Knowler WC, Johannes RS. Using the ADAP Learning Algorithm to Forecast the Onset of Diabetes Mellitus. Proc Annu Symp Comput Appl Med Care. 1988 Nov 9;261–5.

Examples

```
### Loading the PIMA dataset contained in the mlbench package

library(mlbench)
data(PimaIndiansDiabetes)

### Using ELJAlinear to perform EWAS analysis

ELJAlinear(var = 'pregnant',data = PimaIndiansDiabetes,manplot = TRUE,
Bonferroni = TRUE,FDR = TRUE, nbvalmanplot = 30, manplotsign = FALSE)
results
```

`ELJAlogistic` *Logistic regression tool for EnvWAS/EWAS analysis*

Description

A tool for Environment-Wide Association Studies (EnvWAS / EWAS) which are repeated analysis. This function is espacially for logistic regression based on the glm function with a binomial family with a logit link and allows the addition of adjustment variables.

Usage

```
ELJAlogistic(
  var,
  var_adjust = NULL,
  data,
  manplot = TRUE,
  nbvalmanplot = 100,
  Bonferroni = FALSE,
  FDR = FALSE,
  manplotsign = FALSE
)
```

Arguments

`var`	A categorical and binary variable. It is generally your outcome.
`var_adjust`	A vector containing the names of the fixed adjustment variables for all the models.
`data`	A dataframe containing all the variables needed for the analysis.
`manplot`	Generate a Manhattan plot of the results of the analysis.
`nbvalmanplot`	The number of variables to include in each Manhattan plot.
`Bonferroni`	Add a dashed bar to the Manhattan plot showing the Bonferroni significance level.
`FDR`	Add a dashed bar to the Manhattan plot showing the False Discovery Rate (Benjamini-Hochberg method) significance threshold. NA if all p-values > FDR corrected p-values.
`manplotsign`	Generates a Manhattan plot with only significant results (p<0.05).

Value

A Dataframe with results for each variable of the model.

References

Dunn OJ. Multiple Comparisons Among Means. Journal of the American Statistical Association. 1961;56(293):52-64. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289-300. MLBench · Distributed Machine Learning Benchmark. Available from: https://mlbench.github.io/ Smith JW, Everhart JE, Dickson WC, Knowler WC, Johannes RS. Using the ADAP Learning Algorithm to Forecast the Onset of Diabetes Mellitus. Proc Annu Symp Comput Appl Med Care. 1988 Nov 9;261–5.

Examples

```
### Loading the PIMA dataset contained in the mlbench package

library(mlbench)
data(PimaIndiansDiabetes)

### Using ELJAlinear to perform EWAS analysis

ELJAlogistic(var = 'diabetes',data = PimaIndiansDiabetes,manplot = TRUE,
Bonferroni = TRUE,FDR = TRUE, nbvalmanplot = 30, manplotsign = FALSE)
results
```

Index