

Package: Elja (via r-universe)

August 27, 2024

Type Package

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

Version 1.0.0

Maintainer Marwan El Homsı <marwan.el-homsı@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

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

Repository <https://ehmarwan.r-universe.dev>

RemoteUrl <https://github.com/ehmarwan/elja>

RemoteRef HEAD

RemoteSha aca87989e0a4ab4b8b4febd378870d1f971dd325

Contents

| | |
|--------------|---|
| ELJAglm | 2 |
| ELJAlinear | 3 |
| ELJAlogistic | 5 |

| | |
|--------------|----------|
| Index | 7 |
|--------------|----------|

Description

A tool for Environment-Wide Association Studies (EnvWAS / EWAS) which are repeated analysis. This function is especially 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

| | |
|---------------------------|--|
| <code>var</code> | A categorical and binary variable. It is generally your outcome. |
| <code>var_adjust</code> | A vector containing the names of the fixed adjustment variables for all the models. |
| <code>family</code> | The family and the link use for the glm function. |
| <code>data</code> | A dataframe containing all the variables needed for the analysis. |
| <code>manplot</code> | Generate a Manhattan plot of the results of the analysis. |
| <code>nbvalmanplot</code> | The number of variables to include in each Manhattan plot. |
| <code>Bonferroni</code> | Add a dashed bar to the Manhattan plot showing the Bonferroni significance threshold. |
| <code>FDR</code> | 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. |
| <code>manplotsign</code> | 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

| | |
|---------------------------|--|
| <code>var</code> | A categorical and binary variable. It is generally your outcome. |
| <code>var_adjust</code> | A vector containing the names of the fixed adjustment variables for all the models. |
| <code>data</code> | A dataframe containing all the variables needed for the analysis. |
| <code>manplot</code> | Generate a Manhattan plot of the results of the analysis. |
| <code>nbvalmanplot</code> | The number of variables to include in each Manhattan plot. |
| <code>Bonferroni</code> | Add a dashed bar to the Manhattan plot showing the Bonferroni significance level. |
| <code>FDR</code> | 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. |
| <code>manplotsign</code> | 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
```

Description

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

Usage

```
ELJAl Logistic(  
  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

ELJAglm, [2](#)
ELJAlinear, [3](#)
ELJAlogistic, [5](#)