Package 'ExpDesignPkg'

January 18, 2016

Type Package

Title Experimental design

Version 1.1

Date 2015-12-07

Author Georgia Tsiliki

Maintainer Georgia Tsiliki <gtsiliki@central.ntua.gr>

Depends RCurl, pmml, jsonlite, AlgDesign, pls

Description This package performs experimental design by employing algorithmic design R routines from the AlgDesign package. The Brandmaier et al method is also employed.

License GPL-2

NeedsCompilation no

R topics documented:

ExpDesignPkg-package *Experimental design*

Description

Calculates an exact or approximate algorithmic design for one of three criteria, using Federov's exchange algorithm from AlgDesign package

Details

The DESCRIPTION file:

Package:	ExpDesignPkg
Type:	Package
Title:	Experimental design
Version:	1.1
Date:	2015-12-07
Author:	Georgia Tsiliki
Maintainer:	Georgia Tsiliki <gtsiliki@central.ntua.gr>
Depends:	RCurl, pmml, jsonlite, AlgDesign, pls
Description:	This package performs experimental design by employing algorithmic design R routines from the AlgDesign p
License:	GPL-2

Index of help topics:

ExpDesignPkg-package	Experimental design
dat1	A sample data object
dat1i	Information for experimental design function suggest.trials.xy
dat1m	Serialized experimental design model file
dat1p	A sample data object
dat2i	Information for experimental design function suggest.trials.noxy
dat2m	Serialized factorial experimental design model file
exp.design.noxy	Experimental design function for (full) factorial designs
exp.design.xy	Experimental design function with X and/or y values
r2.adj.funct	Adjusted R2 function
r2.funct	R2 function
suggest.trials.noxy	Returns suggested trials for a factorial design
suggest.trials.xy	Returns suggested trials when data are available

The most important functions of the package are exp.design.funct.xy and exp.design.funct.noxy.

They produce a a suggested trials vector/ matrix respectively serialized for suggested.trials.xy and suggested.trials.noxy to handle.

Author(s)

Georgia Tsiliki

Maintainer: Georgia Tsiliki <gtsiliki@central.ntua.gr>

References

Help files of AlgDesign

Examples

```
data("dat1")

predF<- "https://apps.ideaconsult.net/enmtest/property/TOX/UNKNOWN_TOXICITY_SECTION/Net+cell+association/8058C

required.param<- list(nTrials=c(11),criterion='D',form='linear',r2.threshold=0.9)

exp.example<- exp.design.xy(dat1,predF,required.param)
```

---

dat1 *A sample data object*

Description

The dataset for this test is a data frame

Usage

```
data("dat1")
```

Format

A list of two objects

datasetURI a character vector- ambit data set uri

dataEntry a data frame containing two columns: compound and values. Compound is a character vector with all compound ambit uris, and values is a data frame with all numberic values of the protein corona data set (compounds by features). One of the columns is the dependent variable (Net.cell.association) which has some null values- the experimental design algorithm will select some/all of them for next suggested trials.

Details

There are no more details

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1)
## maybe str(dat1) ; plot(dat1) ...
```

`dat1i` *Information for experimental design function suggest.trials.xy*

Description

A list with information for experimental design function suggest.trials.xy

Usage

```
data("dat1i")
```

Format

A list with 8 objects:

`design` a numeric matrix with the suggested design for the data

`selected.rows` the selected rows for all the available combinations of the parameters.

`norm.var` the normalized variance. This is the Ge value from optFederov(): The minimax normalized variance over X, expressed as an efficiency with respect to the optimal approximate theory design. Optimal value is 1.

`confounding.effect` the diagonality of the design, excluding the constant if any, from eval.design() function. The optimal number is 1.

`r.squared` the r2 value for the data supplied

`adj.r.squared` the adjusted r2 value for the data supplied.

`verbal.notes` remarks commenting on Ge and diagonality.

Details

Example dataset to suggest trials together with dat1p, dat1m

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1i)
## maybe str(dat1i) ; plot(dat1i) ...
```

---

dat1m *Serialized experimental design model file*

Description

A character string for a serialized experimental design model, i.e. a list including one vectorindicating which are the next trials that should be conducted.

Usage

```
data("dat1m")
```

Format

A character string

Details

Example experimental design model based on dat1

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1m)
## maybe str(dat1m) ; plot(dat1m) ...
```

dat1p *A sample data object*

Description

The dataset for this test is a data frame

Usage

```
data("dat1p")
```

Format

A list of two objects

datasetURI a character vector- ambit data set uri

dataEntry a data frame containing two columns: compound and values. Exactly the same as dat1.

Details

Data set for prediction with dat1m, although data are not used by the function

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1p)
## maybe str(dat1p) ; plot(dat1p) ...
```

dat2i *Information for experimental design function suggest.trials.noxy*

Description

A list with information for experimental design function suggest.trials.noxy

Usage

```
data("dat2i")
```

Format

A list with 6 objects:

`design` a numeric matrix with the suggested design for the data

`selected.rows` the selected rows for all the available combinations of the parameters.

`norm.var` the normalized variance. This is the Ge value from optFederov(): The minimax normalized variance over X, expressed as an efficiency with respect to the optimal approximate theory design. Optimal value is 1.

`confounding.effect` the diagonality of the design, excluding the constant if any, from eval.design() function. The optimal number is 1.

`verbal.notes` remarks commenting on Ge and diagonality.

Details

Example dataset to suggest trials together with dat2m

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat2i)
## maybe str(dat2i) ; plot(dat2i) ...
```

`dat2m` *Serialized factorial experimental design model file*

Description

A character string for a serialized factorial design, i.e. a list including one vector indicating which are the trials that should be conducted.

Usage

```
data("dat2m")
```

Format

A character string

Details

Example experimental design model produced by exp.design.noxy function

References

There are no references

Examples

```
data(dat2m)
## maybe str(dat2m) ; plot(dat2m) ...
```

`exp.design.noxy` *Experimental design function for (full) factorial designs*

Description

Calculates an exact or approximate algorithmic design for one of three criteria, using Federov's exchange algorithm from AlgDesign package. The user needs to specify the number of variables and their levels, then a design matrix with suggested trials is produced. Please note in this case, no X or Y values are provided.

Usage

```
exp.design.noxy(dataset, predictionFeature, parameters)
```

Arguments

`dataset`	list of 2 objects, datasetURI and dataEntry - NOT required
`predictionFeature`	character string specifying which is the prediction feature in dataEntry - NOT required
`parameters`	A list with 7 objects: levels, nVars (number of variables), factors (which are the factor variables), varNames (variables' names), nTrials (the number of suggested trials, if 0 then an estimated number is suggested), criterion ('A', 'I', 'D'), form of the design ('linear','quad','cubic','cubicS')

Details

No details required

Value

A List

`rawModel` A serialized numeric matrix indicating the experimental design for the various variables and their levels.

`pmmlModel` A pmml object - now empty

`independentFeatures`
A list with the names of the variables as given by the user.

`predictedFeatures`
A character vector with names for the suggested trials

`additionalInfo` A list including the following: design (The design suggested), selected.rows (The rows (nanoparticles) suggested for new trials), norm.values (The minimax normalized variance over X, expressed as an efficiency with respect to the optimal approximate theory design.), confounding.effect (The diagonality of the design, excluding the constant, if any.), verbal.notes (Verbal notes to comment on norm.values and confounding.effect), predictedFeatures (The character string 'suggestedTrials' to indicate the new vector created including all experimental design memberships).

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

The help file of blockcluster package

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--or do  help(data=index)  for the standard data sets.


required.param<-  list(levels=3, nVars=3, factors='null', varNames=c('a','b','c'),nTrials=10,criterion='D',form


exp.example<- exp.design.noxy(null,null,required.param)
```

`exp.design.xy` *Experimental design function with X and/or y values*

Description

Calculates an exact or approximate algorithmic design for one of three criteria, using Federov's exchange algorithm from AlgDesign package

Usage

```
exp.design.xy(dataset, predictionFeature, parameters)
```

Arguments

`dataset`	list of 2 objects, datasetURI:= character sring, code name of dataset, dataEntry:= data frame with 2 columns
`predictionFeature`	character string specifying which is the prediction feature in dataEntry
`parameters`	list with parameter values for experimental design. 4 objects should be included, i.e. nTrials a numeric value indicating number of trials suggested, if 0 then an estimated number is suggested, criterion a character value to indicate which optimal deisgn to apply (possible values are 'D', 'A', 'I'), form a string indicating the formula of the deisgn (possible formulas are 'linear','quad','cubic','cubicS'), r2.threshold a numeric value indicating the r2 threshold value (If the data supplied provides r2 value greater than the threshold value, a warning message is returned.).

Details

No details required

Value

A List

`rawModel`	A serialized numeric vector indicating the experimental design memberships of nanoparticles in the data.
`pmmlModel`	A pmml object - now empty
`independentFeatures`	A list with Ambit names for all genes/ proteins features included in the model
`predictedFeatures`	A character vector with names for the suggested trials
`additionalInfo`	A list including the following: design (The design suggested), selected.rows (The rows (nanoparticles) suggested for new trials), norm.values (The minimax normalized variance over X, expressed as an efficiency with respect to the optimal approximate theory design.), confounding.effect (The diagonality of the design, excluding the constant, if any.), r.squared (when y is given), adj.r.squared

(when y is given), verbal.notes (Verbal notes to comment on norm.values and confounding.effect), predictedFeatures (The character string 'suggestedTrials' and 'newY' to indicate the new vectors created including all experimental design memberships).

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

The help file of blockcluster package

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--  or do  help(data=index)  for the standard data sets.

data("dat1")

predF<- "https://apps.ideaconsult.net/enmtest/property/TOX/UNKNOWN_TOXICITY_SECTION/Net+cell+association/8058C

required.param<- list(nTrials=c(11),criterion='D',form='linear',r2.threshold=0.9)

exp.example<- exp.design.xy(dat1,predF,required.param)
```

`r2.adj.funct` *Adjusted R2 function*

Description

Calculates the adjusted R2 value

Usage

```
r2.adj.funct(y, y.new, num.pred)
```

Arguments

`y`	observed y values
`y.new`	predicted y values
`num.pred`	number of parameters in the predicted model

Details

No details required

Value

A numeric value for the adjusted coefficient of determination R2

Author(s)

Georgia Tsiliki

References

Dobson An introduction to linear modelling

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--or do  help(data=index)  for the standard data sets.

r2.adj.funct(1:10,1:10,2)
```

`r2.funct` *R2 function*

Description

Calculates the R2 value

Usage

```
r2.funct(y, y.new)
```

Arguments

`y`	observed y values
`y.new`	predicted y values

Details

No details required

Value

A numeric value for the coefficient of determination R2

Author(s)

Georgia Tsiliki

References

Dobson An introduction to linear modelling

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--or do  help(data=index)  for the standard data sets.

r2.funct(1:10,1:10)
```

---

suggest.trials.noxy *Returns suggested trials for a factorial design*

Description

Suggested trials are returned as a list with one object, a numeric matrix with the last column indicating with 1 the suggested trial

Usage

```
suggest.trials.noxy(dataset, rawModel, additionalInfo)
```

Arguments

dataset	Data for prediction. A list of two objects: datasetURI (a character string), dataEntry (a data frame).
rawModel	R model serialized (suggested trials for the data matrix supplied in exp.design.funct1).
additionalInfo	Any additional information needed for rawModel. Here the list generated by exp.design.funct.noxy. The list should contain a field named 'predictedFeatures' which should be exactly the same as that returned by exp.design.funct.noxy function.

Details

No further details required

Value

A list of one objected called 'predictions' which is also a list of one cell data-frames each containing the suggested trials for the data tested. This object is a matrix for the suggested trials per variable and level; the last column indicates with 1 the suggested trial

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

No references required

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--or do  help(data=index)  for the standard data sets.


data("dat2m")
data("dat2i")

pred.res<- suggest.trials.noxy(null, dat2m, dat2i)
```

`suggest.trials.xy` *Returns suggested trials when data are available*

Description

Suggested trials are returned as a list with one object, a binary vector where 1 indicates suggested trial

Usage

```
suggest.trials.xy(dataset, rawModel, additionalInfo)
```

Arguments

`dataset`	Data for prediction. A list of two objects: datasetURI (a character string), dataEntry (a data frame).
`rawModel`	R model serialized (suggested trials for the data matrix supplied in exp.design.funct1).
`additionalInfo`	Any additional information needed for rawModel. Here the list generated by exp.design.funct.xy. The list should contain a field named 'predictedFeatures' which should be exactly the same as that returned by exp.design.funct.xy function.

Details

No further details required

Value

A list of one objected called 'predictions' which is also a list of one cell data-frames each containing the suggested trials for the data tested. This object is a binary vector with 1 indicates suggested trial

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

No references required

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--  or do  help(data=index)  for the standard data sets.

data("dat1p")
data("dat1m")
data("dat1i")

pred.res<- suggest.trials.xy(dat1p, dat1m, dat1i)
```

Index