Package 'ares'

February 14, 2012

Version 0.7.2

Date 2011-05-19

Author Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Maintainer Washington Junger <wjunger@ims.uerj.br>

Depends R (>= 2.11.0),utils,stats,foreign,gam,gplots,splines

Title Environment air pollution epidemiology: a library for time series analysis

Description A toolbox for time series analyses using generalized additive models. This library includes routines for model estimation and diagnostics. Polynomial distributed lag models are also implemented.

License GPL (>= 2)

Repository CRAN

Date/Publication 2011-05-20 12:05:08

R topics documented:

admrio *Hospital Admissions in Rio de Janeiro*

Description

This is a typical dataset used in environment air pollution epidemiology. It contains outcome, confounders and exposure variables, i.e., event counts, meteorological condition and air pollutants concentrations.

Usage

```
data(admrio)
```

Format

A data frame with 365 observations on the following 20 variables.

`date` dates string formatted as *dd/mm/yyyy*

`resp65` counts of respiratory hospital admissions over 65 years old

`circ65` counts of circulatory hospital admissions over 65 years old

`copd65` counts of COPD hospital admissions over 65 years old

`pneum65` counts of pneumonia hospital admissions over 65 years old

`myinf65` counts of myocardial infarction hospital admissions over 65 years old

`resp5` counts of respiratory hospital admissions under 5 years old

`pneum5` counts of pneumonia hospital admissions under 5 years old

`warmseason` an indicator variable for months with higher temperature

`tmpmed` mean of measured temperature

`tmpmin` minimum of measured temperature

`tmpmax` maximum of measured temperature

`humid` mean of measured relative humidity

`rain` daily rainfall in *mm*

`rainy` an indicator for days with rainfall over 50 ml

`pm10` daily averaged concentrations of particulate matter up to 10 microns in micrograms per cubic metre

`so2` daily averaged concentrations of sulphur dioxide in micrograms per cubic metre

`co` daily averaged concentrations of carbon monoxide in micrograms per cubic metre

Details

This dataset is a subset of a larger dataset used in an actual study in Rio de Janeiro.

Source

Ares-Rio Program in http://www.ims.uerj.br/ares-rio

References

\ Air quality and effects of the Rio de Janeiro municipality population - Final research report (2005). It can be downloaded from http://www.ims.uerj.br/ares-rio/download/relatorio_ares-rio_mma_jan2005.pdf - Portuguese version only.

Examples

```
data(admrio)
names(admrio)
```

---

bubbleplot *Bubble Plot*

Description

Produce a bubble type plot

Usage

```
bubbleplot(x, y, z, bs=0.1, ...)
```

Arguments

x	a vector of x-axis values
y	a vector of y-axis values
z	a vector of values representing the radius of the bubbles
bs	an expansion factor for the bubbles size
...	further options for plot

Details

This function implements a bubble type plot. One example of its use is plotting some characteristics of the population and pass the weights of each observation as the radius of the bubbles.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

plot

---

count.na *Count Missing*

Description

Count and return some statistics about missing values

Usage

```
count.na(var)
```

Arguments

var	a vector

Details

The function take var and count the occurrence of missing data.

Value

This function returns a list with

n.total	number of observations of var
na	number of missing observations
n.valid	number of valid observations
percent.na	percentage of missing observations

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>
Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

is.na,na.action

Examples

```
data(admrio)
setup(admrio,"date")
count.na(pm10)
```

`daily.stats` *Daily Statistics*

Description

Compute daily basis statistics from hourly measures

Usage

```
daily.stats(dataset, parameter, first.column = 2, date = TRUE,
samples = 24, statistic = "mean", daylight=c(6,19), date.format = "%d/%m/%Y")
```

Arguments

`dataset`	input dataset
`parameter`	quoted string indicating the environmental parameter rules
`first.column`	index of the first column of actual data. The first column is usually allocated for dates, so default is 2
`date`	a logical indicating whether the resulting dataset should contain a date column
`samples`	an integer indicating the number of samples per day
`statistic`	a quoted string indicating the aggregation function. Default is `"mean"`. See Details
`daylight`	a vector indicating the daylight hours. Default is `c(,6,19)`
`date.format`	a string indicating the date format for the horizontal axis. Default is `dd/mm/yyyy`. See `strptime` for options

Details

This is an environmental-specific function. Suppose, for instance, that during the day 24 measurements of a pollutant concentration are taken and one wants to compute daily statistics about the monitored parameter. Besides, when `parameter` is set to some "known" air pollutant, the aggregation is computed according to specific predefined rules.

The available statistics are: `"mean"`, `"min"`, and `"max"`. Whilst *min* and *max* are the trivial maximum and minimum respectively, *mean* are computed using different rules for each pollutant.

The implemented rules for the mean of some pollutants are the following:

`TEMP` or `HUMID`: ordinary daily mean are computed;

`PM10` or `SO2`: daily mean are computed if at least three quarters of hourly observations is not missing;

`NO2` or `O3`: daily mean are computed if at least three quarters of hourly observations is not missing and at least one quarter lies between 6 a.m and 6 p.m.;

`CO`: if parameter is set to `CO`, statistic is ignored and the maximum 8-hour running mean is computed.

Attention! Be sure that all columns on the dataset share the same rule for statistic computation, otherwise it will produce wrong values.

Both `parameter` and `statistic` are not case sensitive.

Value

A data frame with the same columns there are in `dataset` plus the date column.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

See Also

`import.data`,`export.data`

Examples

```
## Not run:
# pm.mean <- daily.stats(pm.filled,"PM10",statistic="mean")
## End (Not run)
```

---

desc.data *Describe Dataset*

Description

Report some information on a dataset

Usage

```
desc.data(dataset = NULL)
```

Arguments

`dataset`	an object of class `data.frame`

Details

This function outputs information on variables index, names, class, number of missing values, and labels. If `dataset` is omitted, the function tries to find `.ares.active.dataset`.

Value

A data frame containing columns with the items described in Details.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

names,class

Examples

```
data(admrio)
setup(admrio,"date")
## specific dataset

## .ares.active.data.set
d <- desc.data()
```

---

desc.vars *Describe Variables*

Description

Compute some common descriptives statistics of variables

Usage

```
desc.vars(vars, by = NULL, dataset = .ares.active.dataset,
stats = c("n", "na", "mean", "sd", "min", "max", "centiles"),
probs = c(0.25, 0.5, 0.75), labels = vars, print = TRUE,
digits = getOption("digits"), ...)
```

Arguments

vars	a vector of quoted strings with the variables names
by	a factor variable for subset analysis.It will be coerced to a factor. Default is NULL
dataset	source data frame. Default is .ares.active.dataset
stats	a vector of quoted strings with the statistics to be computed. See Details
probs	a vector of probabilities for the quantiles computation
labels	a vector of quoted strings with alternate labels for the variables. If omitted vars is used
print	a logical indicating whether the statistics should be printed
digits	an integer indicating the number of decimal places to print. Default is given by the system option digits
...	further arguments for mean, basically, support for trimming

Details

By default, this function computes the following statistics: *n*: number of observations, *na*: number of missing values, *mean*: variable means, *sd*: standard deviation, *min*: minimum value; *max*: maximum value, and *centiles*: quantiles controlled by the vector `probs`. Any subset of these statistics can be selected. However, if other than these is indicated it will throw an error. The statistics in `stats` are not case sensitive.

If the argument by is set, the statistics are computed for each subset defined by the supplied vector.

These statistics are frequently found on publications regarding air pollution and health effects.

Value

The function invisibly returns a data frame with a variable per line and a statistic per column.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Whetherill, G. B. *Elementary Statistical Methods*. 2 ed. Chapman and Hall.

See Also

summary,mean,quantile

Examples

```
data(admrio)
setup(admrio,"date")
## no subset analysis
d <- desc.vars(c("resp5","pm10","so2","co"),stats=c("n","mean","sd","centiles"),digits=3)
## subset analysis
d <- desc.vars(c("resp5","pm10","so2","co"),by=quarters,stats=c("n","mean","sd","centiles"),digits=3)
```

---

diagnostics *Diagnostics*

Description

Perform a series of model dignostics

Usage

```
diagnostics(model, single.graph = TRUE)
```

Arguments

model	a model fitted by fit.core
single.graph	a logical indicating whether all graphs should be plotted on the same device windows

Details

This function is a short-cut to several common diagnostics used when analysing time series of outcomes in environmental epidemiology. The following procedures are executed by diagnostics: plot of fitted values with plot.fitted, plot of residuals with plot.residuals, plot of Cook's distance with plot.cook, plot of partial autocorrelation function with plot.pacf, plot of periodogram with periodogram, quantile-quantile plot with plot.qq, and information on parameters and smooth functions estimates print.summary.

If all the graphs are plotted on the same device window, it can be rather difficult to read, but it is more suitable for publications and the graph can be saved with high quality format using save.plot.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

See Also

fit.core,summary

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
diagnostics(m)
```

---

dispersion *Dispersion Parameter*

Description

Estimate dispersion parameter

Usage

```
dispersion(model)
```

Arguments

model	a model fitted by fit.core

Details

This is an consistent and asymptotic estimation of the dispersion parameter based on generalized Pearson statistics for GLM models. It is given $\frac{\sum_{t=1}^{n}(y_t - E(y_t))^2/E(y_t)}{n-df}$, where n is the number of observations and df is the number of degrees of freedom used by the model.

Value

A scalar with the statistic.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

fit.core,gam,glm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
dispersion(m)
```

`estimate.risks` *Estimate Risks*

Description

Estimate the effects of the pollutants

Usage

```
estimate.risks(model, pollutant, unit = rep(10,length(pollutant)), confidence.level = .95, method = "si
perc.rr = TRUE, interaction = NULL, lag.struc = list(l = 0:5, ma = NULL), pdlm.struc = list(l = 5,d = rep(
length(pollutant)), overall = TRUE), overdispersion = FALSE, labels = pollutant, print = TRUE, digits =
getOption("digits"), plot = TRUE, new = TRUE, graph.scale = FALSE, verbose = TRUE, ...)
```

Arguments

`model`	a model fitted by `fit.core`
`pollutant`	a vector with the names of he variables to estimate the effects
`unit`	a vector indicating the units for relative risk computation. Default is 10 for all pollutants. See Details
`confidence.level`	confidence level for interval computation
`method`	a string indicating the method for effect estimation. Default is `"singlelag"`. See Details
`perc.rr`	logical. If `TRUE` the effects are reported in terms of percent changes in risk. Default is `TRUE`
`interaction`	a string indicating the 2-level interaction term usually used for seasonal effects estimation. Default is `NULL`. See Details
`lag.struc`	a list with the single lag model structure. Default is lag up to 5 days. See Details
`pdlm.struc`	a list with the polynomial distributed lag model structure. Default is lag up to 5 and a 2-degree polynomial function. If `overall=TRUE` overall effect is computed. See Details
`overdispersion`	a logical indicating whether confidence intervals should account for the extra variability
`labels`	a vector of quoted strings with alternate labels for the pollutants. Default is the names of the variables in `pollutant`
`print`	a logical indicating whether the statistics should be printed

digits	an integer indicating the number of decimal places to print. Default is given by the system option digits
plot	a logical indicating whether the estimated risks should be plotted. Default is TRUE. See plot.risk
new	if TRUE a new graph window is opened
graph.scale	can be either a logical or a vector with the axis limits. If TRUE or a vector all the graphs will share the same y-axis scale
verbose	a logical indicating whether extra information should be printed during the iterations
...	further options for plot.risk

Details

This function estimates the effects for each pollutant in pollutant using the estimation approach set in method. If method is set to either "singlelag" or "dual" the effects are estimated independently for each exposure. If it is set to "pdlm" then a polynomial distributed lag model is used to estimate the effects using the lag structure passed to pdlm using pdlm.struc option.

The lag.struc argument is a list containing the number of lags (l) and/or the moving averages (ma) of the pollutants as the measure of exposure. This mode allows more than one pollutant at a time. The general list structure is list(l=,ma=,ma.base=,labels=), where l and ma are vectors indicating the lagged exposures in order to estimate the effects. If ma.base is omitted or ma.base=0, the moving averages will range from the concurrent day to each element in ma. If labels is missing or set to NULL, a generic label will be used.

The pdlm.struc argument is a list containing the number of lags (l) and the degrees (deg) to be passed to pdlm. This mode allows more than one pollutant at a time. The general list structure is list(l=,deg=,labels=), where l is an integer and deg is an integer or vector with the same length as pollutant. If deg is an integer all the pollutants will share the same polynomial structure. If labels is missing or set to NULL, a generic label will be used instead. Overall estimate and confidence interval will be plotted if overall=TRUE.

Dual pollutant models may be estimated by setting the option method to "dual". This method will estimate a model for each combination of two of the pollutants set in pollutant. One can set the option lag.struc the same way as in the single lag models, however both pollutants in the model will share the same lag structure, i.e., the effects of both pollutant will be estimated by using the same lagged exposure. It is not a serious limitation though. Due to the manner the pollutants effects matrices are stored in the risk array, one should avoid reading them directly. Use the print.risk function instead.

If plot is set to TRUE and the method is dual pollutant models, then the user will have to choose which pollutant should be plotted. It is not possible to plot all of them at once. A handy menu is provided for selection.

A 2-level interaction term can be supplied in interaction. The interaction effect will be estimated as well as the marginal effects. The argument interaction must be a factor or it will be coerced to a factor. Interaction estimation is available for single lag method only.

Over-dispersed models can be fitted by *quasi-likelihood* if overdispersion=TRUE. One should get larger confidence intervals when this options is set under over-dispersed data.

Value

The function invisibly returns an array of matrices with the exposures on the rows ant the statistics on the columns.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

fit.core,print.risk,plot.risk

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
## single lag effect estimation
r1 <- estimate.risks(m,c("pm10","so2"),digits=3,labels=c("PM10","SO2"),method="singlelag",lag.struc=list(l=0:2,

## pdlm effect estimation
r2 <- estimate.risks(m,c("pm10","so2"),digits=3,labels=c("PM10","SO2"),method="pdlm",pdlm.struc=list(l=5,deg=c(

## dual pollutant model (it is commented to not run during check)
## r3 <- estimate.risks(m,c("pm10","so2","co"),digits=3,labels=c("PM10","SO2","CO"),method="dual",lag.struc=lis
```

---

explore.humid *Explore Humidity*

Description

Explore the exposure-response relationship of humidity

Usage

```
explore.humid(model, var, df = 4, type = "deviance", new = TRUE, ...)
```

Arguments

model	a model fitted by fit.core
var	the humidity variable
df	degrees of freedom for the smoother. If set to 0 the smooth line is not plotted
type	a quoted string with the type of residuals to plot
new	if TRUE a new graph window is opened
...	further options for plot

Details

The scatter plot of residuals against humidity is smoothed by means of a spline function defined by df.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

plot,smooth.spline,setup

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)
m <- fit.core(f)
explore.humid(m,humid,df=5)
```

`explore.temp` *Explore Temperature*

Description

Explore the exposure-response relationship of temperature

Usage

```
explore.temp(model, var, df = 4, type = "deviance", new = TRUE, ...)
```

Arguments

`model`	a model fitted by `fit.core`
`var`	the temperature variable
`df`	degrees of freedom for the smoother. If set to 0 the smooth line is not plotted
`type`	a quoted string with the type of residuals to plot
`new`	if `TRUE` a new graph window is opened
`...`	further options for `plot`

Details

The scatter plot of residuals against temperature is smoothed by means of a spline function defined by `df`.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

`plot`,`smooth.spline`,`setup`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(humid)
m <- fit.core(f)
explore.temp(m,tmpmax,df=7)
```

`export.data` *Export Data*

Description

Export datasets into different formats

Usage

```
export.data(data, file, text.format = "csv", ...)
```

Arguments

`data`	a matrix or data frame
`file`	quoted string with the file name
`text.format`	rules for text files
`...`	further options for `write.table`

Details

The function will try to select the proper filter based on the file extension. Valid file types are those supported by the package **foreign**. There are rules implemented for Stata up to version 9 (*.dta) and xBase files up to version 4 (*.dbf). Excel files (*.xls) are no longer supported. R binary files (*.rda) can also be written directly by the function.

Several text formats can be written by setting `text.format` or passing options through `...`. If none of the above proprietary formats extension is part of the file name, the function will save the data frame in ASCII text format disregarding the extension. Available options for `text.format` are `"csv"` for comma separated values, `"tab"` for columns separated by a tab character and `"spc"` for space separated columns and, also, its variants for comma as the decimal separator, `"csv2"`, `"tab2"` and `"spc2"`.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

attach,search

Examples

```
data(admrio)
export.data(amdrio,"amdrio.dta")
```

---

exposure.response *Explore Exposure-Response Curve*

Description

Explore the exposure-response relationship of the pollutant

Usage

```
exposure.response(model, var, df = 4, type = "deviance", new = TRUE, ...)
```

Arguments

model	a model fitted by fit.core
var	the temperature variable
df	degrees of freedom for the smoother. If set to 0 the smooth line is not plotted
type	a quoted string with the type of residuals to plot
new	if TRUE a new graph window is opened
...	further options for plot

Details

The scatter plot of residuals against the pollutant concentrations is smoothed by means of a spline function defined by df.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

plot,smooth.spline,setup

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
exposure.response(m,pm10,df=7)
```

`fillup.hours` *Fill up Hours*

Description

Fill up missing hours in the hourly measured datasets

Usage

```
fillup.hours(input, date.time, readings = 24, cycle = seq(0, 23, by = 24/readings),
start.date = NULL, end.date = NULL, date.format = "%d/%m/%Y", hour.format =
"%H:%M:%S", offset = 0, sep = " ")
```

Arguments

`input`	a data frame. See Details
`date.time`	a quoted string with the name of a date and time formatted variable
`readings`	number of measures taken a day. Default is 24 (hourly)
`cycle`	a vector with the hours measures are taken. Default is a sequence from 0 to 24/`measures`
`start.date`	initial date for the dummy dataset
`end.date`	final date for the dummy dataset
`date.format`	date format. See `strptime` for options
`hour.format`	hour format. See `strptime` for options
`offset`	an integer with the number of minutes to shift minutes to `"00"`. Default is 0
`sep`	character to separate date and time formats. Default is a space character

Details

This is an environmental-specific function. Suppose, for instance, that during the day, 24 measurements of a pollutant concentration are taken and if it is missing, the monitor keep the entry out of the log. Thus, at the end of the day less than 24 hourly entries are recorded. However, one needs a regular 24-hour vector to carry on with daily statistics computations, i.e., means, quantiles, etc.

This function reads the input dataset from a data frame. A complete dummy dataset with NA entries is built from the input parameters. Finally, the original dataset and the dummy dataset are merged using the date.time variables as reference. The resulting filled up dataset may be output to a data frame or directly to a file which format is supported by export.data.

Value

This function returns a data frame with missing filled up with NA.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

import.data,export.data

Examples

```
## Not run:
# pm.filled <- fillup.hours(pm_hourly,"Date.Time")
## End (Not run)
```

---

fit.core *Fit Core Model*

Description

Fit the core model

Usage

```
fit.core(formula, weights = NULL, class = "gam", ...)
```

Arguments

formula	the formula for the model. See formula
weights	an optional vector of weights to be used in fitting process. See Details
class	a class for the model to be used in the fitting process. Default is gam. See Details
...	further arguments for the model fitting function

Details

The argument `weights` is an optional vector for weighting. Usually, it is a vector of ones if none is specified, but when the user has prior information about the weights of each observations, it can be set accordingly. It is particularly useful when the dataset has suffered some sort of imputation technique.

This function is primarily a wrapper for the real fitting function. For now, `gam` and `glm` are the allowed classes. However, other fitting methods can be added in the future. Some information is added to the resulting object and therefore the class `ares` is added to its inheritance. For details on GAM options and output see `gam` for GLM options and output see `glm`.

Value

An object of the same class as the one specified by `class`. See each class help for a list of objects returned.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

`plot`,`smooth.spline`,`setup`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
print.summary(m)
```

gen.holidays *Generate Holidays*

Description

Generate indicator variables for holidays and other special dates

Usage

```
gen.holidays(date, holidays = NULL, dates = NULL,
selection = TRUE, country = NULL)
```

Arguments

date	date formatted vector. See as.Date
holidays	vector of strings for the holidays names
dates	vector of strings for the holidays dates formatted as "dd/mm" or "dd/mm/yyyy". See Details
selection	a logical indicating whether .ares.selection should honoured
country	a string indicating the country rules. Default is NULL. See Details

Details

This function scans date looking for matches to each element of dates and if it finds any 1 is assigned to the variable 0 otherwise. The column is named according to holidays. If both holidays and date are omitted and country is set to a valid rule, country-specific holidays are included in the matrix. If both .holidays and .dates vectors are set in the global environment they are used instead of country rules (that are overridden) to flag the fixed holidays. Moving holidays based on Easter are computed by the function moving.holidays.

Automatically created holidays are Christmas, New year, Carnaval, Passion, Easter and Corpus Christi. If both holidays and date are supplied by the function caller then the resulting matrix will contain only those holidays and these automatically created holidays will be omitted. Thus, one can use this function for both automatically create holidays and/or some special days indicators. The vector dates can contain dates formatted as both "dd/mm" and "dd/mm/yyyy". The first format will flag all occurrences of that day and month in date and the second, obviously, will flag the only one occurrence.

Other country rules can be added to this function in the future. If you have a list of holidays for you country, send us a note.

Value

A matrix of indicator variables for the holidays with as many rows as date.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

See Also

attach,search,as.Date

Examples

```
data(admrio)
names(admrio)
setup(admrio,"date")
# some mexican holidays plus moving holidays
.holidays <- c("constitucion","cincomayo","independencia")
.dates <- c("05/02","05/05","16/09")
mx.holidays <- gen.holidays(doe)

# special days
strikes <- gen.holidays(doe,holidays=c("transportation","hospitals"),dates=c("02/04/2001","21/05/2001"))
```

---

get.beta *Get Coefficient*

Description

Transform the constrained coefficients back to the unconstrained form

Usage

```
get.beta(coeff, var.coeff, lags, degrees, prefix = "beta")
```

Arguments

coeff	a vector of all the coefficients from the distributed lag model
var.coeff	a matrix of variances and covariances of the coefficients from the distributed lag model
lags	an integer indicating the lags used for the distributed lag structure
degrees	an integer indicating the degrees used for the distributed lag structure
prefix	a quoted string to prefix the coefficients names

Details

This function is used internally by pdlm to put the constrained polynomial distributed lag model coefficients, say η, back to its original scale β.

Value

The function returns a list containing

beta	the unconstrained coefficients
se.beta	the unconstrained coefficients standard errors

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

pdlm

---

get.residuals *Get Residuals*

Description

Extract adjusted residuals from the model

Usage

```
get.residuals(model, type = "adj_deviance", plot = FALSE, ...)
```

Arguments

model	a model fitted by fit.core
type	a quoted string indicating which type of residual to extract. Default is "adj_deviance". See Details
plot	a logical indicating whether the residuals should be plotted. See plot.residuals
...	further options for residuals

Details

The argument `type` may be either `"deviance"`, `"std_deviance"`, `"std_scl_deviance"` or `"adj_deviance"`. Each of them behaves as described bellow.

deviance: Deviance residuals are estimated by $r_t = sign(y_t - E(y_t)) * \sqrt(d_t)$, where d_t is the deviance contribution of the t-th observation. See `deviance` for details on deviance component extraction.

std_deviance: The deviance component is divided by $(1 - h_t)$, where h_t is the t-th element of the diagonal of the pseudo hat matrix of the approximating linear model. So they turn into $r_t = sign(y_t - E(y_t)) * \sqrt(d_t/(1 - h_t))$.

std_scl_deviance: Just like the last one except for the dispersion parameter in its expression, so they have the form $r_t = sign(y_t - E(y_t)) * \sqrt(d_t/\phi * (1 - h_t))$, where ϕ is the estimated dispersion parameter of the model. See `dispersion` for ϕ estimation.

adj_deviance: These are the deviance residuals multiplied by the estimated coefficient of skewness of the distribution. Thus, for a Poisson model they take the form $r_t = sign(y_t - E(y_t)) * \sqrt(d_t) * K_t$, where K_t is given by $K_t = 1/(6\sqrt(E(y_t))$.

Pierce and Shafer (1986) propose the use of the adjusted deviance residuals over other type of residuals.

Value

A vector of class `residuals` with extracted and adjusted residuals of the model.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

See Also

`get.residuals`,`gam`,`glm`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
r <- get.residuals(m,"adj_deviance")
## plot using appropriate method
plot(r)
```

`import.data` *Import Data*

Description

Import datasets in different formats

Usage

```
import.data(file, text.format = "csv", ...)
```

Arguments

`file`	quoted string with the file name
`text.format`	rules for text files
`...`	further options for `read.table`

Details

The function will try to select the proper filter based on the file extension. Valid file types are some of those supported by package **foreign**. There are rules implemented for Stata up to version 9 (*.dta), SPSS up to version 11 (confirmed) (*.sav), xBase files up to version 4 (*.dbf), Epi-Info files version 6 (*.rec). R binary files `*.rda` can also be read directly by this function. Excel files (*.xls) are no longer supported. Please, save the files in the `CSV` format.

Several text formats can be read by setting `text.format` or passing options through `...`. If none of the above proprietary formats extension is part of the file name, the function will assume an ASCII text filter disregarding the extension. Available options for `text.format` are `"csv"` for comma separated values, `"tab"` for columns separated by a tab character and `"spc"` for space separated columns and, also, its variants for comma as the decimal separator, `"csv2"`, `"tab2"` and `"spc2"`.

If `file` is set to `"clipboard"` the function will try to import data previously copied into the system clipboard using text data rules set in `text.format`. It can render unexpected results though.

Value

An object of class `data.frame`.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

attach,search

Examples

```
data(admrio)
export.data(admrio,"admrio.dta")
new.data=import.data("admrio.dta")
```

`l` *Lagged Variables*

Description

Compute lagged variables

Usage

```
l(var, k, selection = TRUE)
```

Arguments

`var`	variable for lagging
`k`	lag
`selection`	a logical indicating whether `.ares.selection` should honoured

Details

This function is used mainly to assess lagged exposure to meteorological factors and/or pollutants. The argument `k` controls the lag distance.

The function may be used inside a model formula to estimate the lagged exposure effects.

Value

Return the variable `var` lagged. If `.ares.selection` is set, lagged `var` is cut accordingly.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

mean,setup

Examples

```
data(admrio)
setup(admrio,"date")
# compute lag2 of pm10
pm10.l2 <- l(pm10,2)
```

`lagvar` *Matrix of lagged variable*

Description

Create a matrix of a variable lagged up to some instant

Usage

```
lagvar(var, k)
```

Arguments

`var`	a vector of the variable to lag
`k`	an integer indicating the number of lags to compute

Details

This function creates a matrix with $k + 1$ columns. The columns hold `var` lagged from 0 to k.

It is designed to be called from within `pdl`.

Value

A matrix with the lagged variate.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

Diggle, P. J. (1990) *Time Series: A Biostatistical Introduction*. Oxford University Press.

See Also

`pdlm`

`ljungbox.test` *Ljung-Box Test*

Description

Run the Ljung-Box test

Usage

```
ljungbox.test(x, k = 25, ...)
```

Arguments

`x`	a vector, usually of residuals
`k`	an integer indicating the number of lags to compute the autocorrelation coefficients
`...`	further options for `acf`

Details

This function implements the the Ljung-Box test for autocorrelation coefficients. The null hypothesis of the test is $H_0 : \rho(1) = \rho(2) = \cdots = \rho(k)$ for $k < (n-1)$. The test statistic depends on `k`.

Value

A list containing the following objects

`statistic`	the test statistic
`p.value`	the p-value
`df`	the degrees of freedom of the chi-squared distribution of the test statistic
`n.used`	the number of observations used for autocorrelations computation
`data.name`	the name of the input variable

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Spanos, A. (1999) *Probability Theory and Statistical Inference*. Cambridge.

See Also

get.residuals,acf

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)
m <- fit.core(f)
r <- get.residuals(m)
ljungbox.test(r,25)
```

---

lspline *Linear Spline*

Description

Generate a basis for a linear spline for piecewise regression

Usage

```
lspline(var, knots = NULL, nknots = NULL, percentiles = NULL, marginal = FALSE, names = NULL)
```

Arguments

var	the covariate to generate the basis
knots	a vector indicating the positions of the knots. See Details
nknots	an integer indicating the number of knots. See Details
percentiles	a vector of percentiles indicating the positions of the knots. See Details
marginal	logical indicating whether the effect for each interval of the spline represents the change in slope from the preceding one. Default is FALSE, so the coefficients represent the slope of the interval
names	a vector of quoted strings with alternate names for the columns of the resulting matrix. The length must be the number of knots plus 1

Details

This functions implements in R the Stata mkspline command. For further references see the article Gould (1993).

It is necessary to set either of knots, nknots, or percentiles. If more than one is specified, it will throw an error. If nknots is specified the intervals created will be equally spaced.

The function is designed to be called directly from within a formula. Therefore, there is no need to create the basis prior to fitting the model.

This function is very useful for exploring the exposure-response curve of the pollutants.

Value

This function returns a matrix with the basis for a linear spline. The number of columns in the matrix is the number of knots plus 1. The columns are named accordingly or after `names`

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Gould, W. W. (1993) Linear splines and piecewise linear functions. *Stata Technical Bulletin* **15**, 13–17.

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

See Also

`smooth.spline`,`setup`

Examples

```
data(admrio)
setup(admrio,"date")
sp.pm10.1 <- lspline(pm10,knots=c(52.50,61.02,74.75))
sp.pm10.2 <- lspline(pm10,nknots=3)
sp.pm10.3 <- lspline(pm10,percentiles=c(0.25,0.50,0.75))

## within a formula
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)+lspline(pm10,knots=c(52.50,61.02,74.75))
m <- fit.core(f)
print.summary(m)
```

---

ma *Moving Average*

Description

Compute simple moving average of a variable.

Usage

```
ma(var, begin, end, selection = TRUE)
```

Arguments

var	variable to average
begin	initial index
end	final index
selection	a logical indicating whether .ares.selection should honoured

Details

This function is used mainly to compute accumulated exposure to meteorological factors and/or pollutants. The arguments begin and end control the window size for simple moving average computation, for instance, if one sets this parameters to 0 and 2 respectively, the resulting averages will be based on the concurrent day and the two previous days.

The function may be used inside a model formula to estimate the accumulated exposure effects.

Value

Return var averaged and if .ares.selection is set, averaged var is cut accordingly.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

mean,setup

Examples

```
data(admrio)
setup(admrio,"date")
# compute moving average of concurrent and previous two days
pm10.ma02 <- ma(pm10,0,2)
```

---

moving.holidays *Moving Holidays*

Description

Compute moving holidays based on Easter

Usage

```
moving.holidays(year)
```

Arguments

year	a numeric or character vector of 4-digits years. It will be coerced to numeric

Details

Ecclesiastical holidays like Easter, Carnaval, Passion and Corpus Christi are efficiently computed.

Value

This function returns a list of vectors with the same length as number of years for each date described in Details.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Astronomical Society of South Australia at http://www.assa.org.au/edm.html

See Also

attach,search,as.Date

---

pdf.report *Output a report to PDF*

Description

Output the analysis results and diagnostics to a PDF file

Usage

```
pdf.report(model, file, pollutants, method = "pdlm", labels = toupper(pollutants),
unit=10, outcome.label = NULL, city = NULL, df = 0, ...)
```

Arguments

model	a model fitted by fit.core
file	a string with a file name to save the output, i.e., "output.pdf"
pollutants	a vector with the names of the variables to estimate the effects
method	estimation method to be used to estimate the pollutant effect. Default is pdlm for distributed lag models. Use both for both single lag and PDLM estimates
labels	a vector of quoted strings with alternate labels for the pollutants. Default is the names of the variables in pollutant

unit	a vector indicating the units for relative risk computation. Default is 10 for all pollutants. See Details
outcome.label	an alternate label for the outcome variable
city	a string indicating the city. It is for the header
df	the number of degrees of freedom for the outcome variable smoothing. Default is 0 for no smoothing
...	further options for estimate.risks

Details

This function outputs several diagnostics plots and statistics of the core model. It also outputs the effects estimates. The results are piped to a PDF device. See pdf for details.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

fit.core,print.risk,plot.risk

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)

## pdlm effect estimation
pdf.report(m,"test.pdf",pollutants=c("so2","pm10"),labels=c("SO2","PM10"),unit=10)
```

pdl *Polynomial Distributed Lag Basis*

Description

Set up a basis for a constrained polynomial distributed lag model

Usage

```
pdl(var, lags, degrees)
```

Arguments

var	a vector with the variate to constrain on the polynomial functional
lags	an integer indicating the maximum lags for the basis
degrees	an integer indicating the degrees used for the polynomial in the distributed lag structure

Details

Each column of the resulting matrix but the first one will be the form $W_d = Z_1 + 2^d Z^2 + .. + q^d Z_q$ and the first one $W_0 = Z_0 + Z_1 + .. + Z_q$, where Z_i is var at lag i.

This function is intended to be called internally by pdlm.

Value

This function returns a matrix with degrees+1 columns.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

pdlm

pdlm *Polynomial Distributed Lag Models*

Description

Fit a constrained polynomial distributed lag model

Usage

```
pdlm(model, var, lags = 5, degrees = 2, ...)
```

Arguments

model	a model fitted by fit.core
var	a vector with the exposure variate or a quoted string with its name
lags	an integer indicating the number of lags to estimate the effects. Default is 5
degrees	an integer indicating the number of degrees for the constrained polynomial. Default is 2 for a parabolic shape
...	arguments passed on to other methods. See family

Details

This function updates model with the unconstrained distributed lag models using pdl. Then, the unconstrained coefficients and their standard errors are extracted using get.beta.

This model is thoroughly discussed in Schwartz (2000).

Value

The class pdlm is added to the model inheritance and the following list is returned

cmodel	the fitted constrained model
variate	the vector with exposure variate data
var.name	the name of exposure variate
beta	the unconstrained coefficients
lags	an integer indicating the lags used for the distributed lag structure
degrees	an integer indicating the degrees used for the polynomial in the distributed lag structure
call	function call

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

gam,glm,fit.core

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
dlm <- pdlm(m,pm10,lags=5,degrees=2)
```

---

periodogram *Periodogram*

Description

Compute and plot the periodogram of the residuals

Usage

```
periodogram(x, type = "deviance", print = TRUE, rows = 20,
test = TRUE, new = TRUE, digits = getOption("digits"))
```

Arguments

x	a model fitted by fit.core or a vector of residuals
type	a quoted string indicating the type of residuals to extract. Default is "deviance". See get.residuals
print	a logical indicating whether the statistics should be printed
digits	an integer indicating the number of decimal places to print. Default is given by the system option digits
rows	an integer giving the number of rows to print
test	a logical indicating whether the uniformity test should be run. See periodogram.test
new	if TRUE a new graph window is opened

Details

This function computes the Fourier frequencies $\omega = 2\pi j/n$ with $j < n/2$ and the periodogram intensity $I(\omega)$. A graph of $I(\omega)$ against the frequencies is plotted.

If object is a model then the residuals are extracted otherwise the periodogram is computed for the vector supplied.

Value

A data frame with the frequencies, periods, and intensities sorted by the latest in decreasing order.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Box, G., Jenkins, G., Reinsel, G. (1994) *Time Series Analysis : Forecasting and Control*. 3 ed., Prentice Hall.

Diggle, P. J. (1990) *Time Series: A Biostatistical Introduction*. Oxford University Press.

See Also

fit.core,get.residuals

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
periodogram(m)
```

---

periodogram.test *Periodogram Test*

Description

Uniformity test for the periodogram

Usage

```
periodogram.test(object, plot = TRUE)
```

Arguments

object	an object created by periodogram
plot	a logical indicating whether the cumulative probability should be plotted

Details

A Kolmogorov-Smirnov (KS) test is performed to evaluate de uniformity of the $I(\omega)$.

See `ks.test` for details on KS tests.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Box, G., Jenkins, G., Reinsel, G. (1994) *Time Series Analysis : Forecasting and Control*. 3 ed., Prentice Hall.

Diggle, P. J. (1990) *Time Series: A Biostatistical Introduction*. Oxford University Press.

See Also

`fit.core`,`get.residuals`,`periodogram`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
p <-periodogram(m,test=FALSE)
periodogram.test(p)
```

---

pgps *Generalized Pearson's Statistic*

Description

Compute Generalized Pearson's statistic for a Poisson GAM model

Usage

```
pgps(model)
```

Arguments

`model`	a model fitted by `fit.core`

Details

This function extracts necessary information from `model` and compute $\sum_{t=1}^{n} (y_t - E(y_t))^2 / E(y_t)$.

Value

A scalar with the statistic.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

`fit.core`,`gam`,`glm`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
pgps(m)
```

---

plot.constinfo *Plot Constant Information*

Description

Plot constant information on the response scale

Usage

```
plot.constinfo(x, type = "deviance", new = TRUE, ...)
```

Arguments

`x`	a model fitted by `fit.core`
`type`	a quoted string indicating the type of residuals to extract. Default is `"deviance"`. See `get.residuals`
`new`	if `TRUE` a new graph window is opened
`...`	further options for `plot`

Details

This function extracts the residuals of x and generates a plot of the residuals against 2 times the square root of the fitted values.

Value

This function does not resturn a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Cook, R. D., Weisberg, S. (1982) *Residuals and Influence in Regression*. Chapman and Hall.

Atkinson, A.C. (1985) *Plots, Transformations and Regression*. Oxford University Press.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

See Also

fit.core,get.residuals

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
plot.constinfo(m)
```

---

plot.cook *Plot Cook*

Description

Plot Cook's distance

Usage

```
plot.cook(x, type = "deviance", line = 0.1, new = TRUE, ...)
```

Arguments

x	a model fitted by fit.core
type	a quoted string indicating the type of residuals to extract. Default is "deviance". See get.residuals
line	a value indicating where to plot a horizontal line on the y-axis. Set NULL for none
new	if TRUE a new graph window is opened
...	further options for plot

Details

This function extracts the residuals of x, computes de Cook's distance, and generates a plot for each observation. It is a very useful tool for influence analysis.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Cook, R. D., Weisberg, S. (1982) *Residuals and Influence in Regression*. Chapman and Hall.

Atkinson, A.C. (1985) *Plots, Transformations and Regression*. Oxford University Press.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

See Also

fit.core,get.residuals

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
plot.cook(m)
```

`plot.envelope` *Plot Envelope*

Description

Plot a simulated envelope of the residuals

Usage

```
plot.envelope(x, rep = 19, type = "deviance", new = TRUE, ...)
```

Arguments

`x`	a model fitted by `fit.core`
`rep`	an integer indicating the number of replications. Default is 19. See Details
`type`	a quoted string indicating the type of residuals to extract. Default is `"deviance"`. See `get.residuals`
`new`	if `TRUE` a new graph window is opened
`...`	further options for `qqnorm`

Details

This function implements a simulated envelope of residuals. It is a plot to check if the normality assumptions of the residuals hold. Although the authors originally suggested 19 replications, at least a hundred may be reasonable, since the processing power of modern computers is no longer a limitation. However, the re-sampling strategy of the method is still time consuming.

The algorithm is implemented for Poisson family models only. Although it supports *quasi-poisson* models, the results may be inaccurate.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>
Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Cook, R. D., Weisberg, S. (1982) *Residuals and Influence in Regression*. Chapman and Hall.

Atkinson, A.C. (1985) *Plots, Transformations and Regression*. Oxford University Press.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

See Also

fit.core,get.residuals

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
plot.envelope(m)
```

---

plot.event *Plot Event*

Description

Plot the daily series of event counts.

Usage

```
plot.event(x, df = 4, gaps = FALSE, type = "p",
title = NULL, date.format = "%d/%m/%Y", new = TRUE, ...)
```

Arguments

x	a vector of daily counts
df	an integer with the degrees of freedom of the spline. If set to 0, no line is plotted
gaps	a logical indicating whether .ares.selection should be honoured
type	a string indicating the type of graphic. Use "p" for points and "l" for line. Default is "p". See plot
title	a string supplying a title for the graphic. If NULL, an automatic title is generated
date.format	a string indicating the date format for the horizontal axis. Default is dd/mm/yyyy. See strptime for more options
new	if TRUE a new graph window is opened
...	further options for plot

Details

This function and plot.pollutant are essentially the same, only some default options differ. There are two of them for convenience only.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Peng, R., Dominici, F. (2008) Statistical methods for environmental epidemiology with R.

See Also

plot,smooth.spline,setup

Examples

```
data(admrio)
setup(admrio,"date")
plot.event(resp5,df=7)
```

`plot.fitted` *Plot Fitted*

Description

Plot the observed and predicted values

Usage

```
plot.fitted(x, gaps = FALSE, date.format = "%d/%m/%Y", new = TRUE, ...)
```

Arguments

x	a model fitted by fit.core
gaps	a logical indicating whether .ares.selection should be honoured
date.format	a string indicating the date format for the horizontal axis. Default is dd/mm/yyyy. See strptime for more options
new	if TRUE a new graph window is opened
...	further options for plot

Details

This function produces a plot with predicted values (points) and fitted ones (line).

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>
Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.
Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

get.residuals,gam,glm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
plot.fitted(m)
```

---

plot.pacf *Plot Partial Autocorrelation Function*

Description

Plot the autocorrelation function of the residuals

Usage

```
plot.pacf(x, lags = 25, acf.too = FALSE, type = "deviance", new = TRUE, ...)
```

Arguments

x	a model fitted by fit.core
lags	an integer indicating the number of lags to compute the partial autocorrelation coefficients
acf.too	a logical indicate whether the autocorrelation coefficients should be computed too. Default is FALSE
type	a quoted string indicating the type of residuals to extract. Default is "deviance". See get.residuals
new	if TRUE a new graph window is opened
...	further options for pacf

Details

This function computes the partial autocorrelation function (ACF) from the residuals `x` and plot it up to `lags`. This function was first designed to plot partial autocorrelation coefficients only. However, as some users prefer to analyse de autocorrelation coefficient, the latest was added as option `acf.too`.

Value

The function returns a list with the autocorrelations coefficients up to `lags`

`acf`	a vector of autocorrelations coefficients. If `acf.too` is set to `FALSE` it is set to `NULL`
`pacf`	a vector of partial autocorrelations coefficients

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Box, G., Jenkins, G., Reinsel, G. (1994) *Time Series Analysis : Forecasting and Control*. 3 ed., Prentice Hall.

Diggle, P. J. (1990) *Time Series: A Biostatistical Introduction*. Oxford University Press.

See Also

`fit.core`,`get.residuals`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
plot.pacf(m)
```

---

plot.pdlm *Plot of PDL Models*

Description

Plot the effects estimated with a polynomial distributed lag model

Usage

```
## S3 method for class 'pdlm'
plot(x, unit = 10, confidence.level = 0.95,
labels = NULL, new = TRUE, ...)
```

Arguments

x	a model of class pdlm created by pdlm
unit	a vector indicating the units for relative risk computation. Default is 10. See estimate.risks for more options
confidence.level	confidence level for interval computation
labels	a vector of quoted strings with alternate labels for the pollutants. Default is the time lag index. See Details
new	if TRUE a new graph window is opened
...	further options for stockplot

Details

This function extracts the coefficients and their standard errors from a polynomial distributed lag model in x and plots a proper graph.

If labels is a vector, it must have the same length as the number of lags plus 1; if it is a quoted string it will be used as a prefix for the labels with the time lag index; and if it is NULL automatically created labels "Lag i" will be used.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>
Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

pdlm,estimate.risks,plot.risk

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
dlm <- pdlm(m,pm10,lags=5,degrees=2)
plot(dlm)
```

`plot.pollutant` *Plot Pollutant*

Description

Plot the daily series of contaminant concentrations.

Usage

```
plot.pollutant(x, df = 4, gaps = FALSE, type = "l",
title = NULL, date.format = "%d/%m/%Y", new = TRUE, ...)
```

Arguments

`x`	a vector of concentrations
`df`	an integer with the degrees of freedom of the spline. If set to 0, no line is plotted
`gaps`	a logical indicating whether `.ares.selection` should be honoured
`type`	a string indicating the type of graphic. Use `"p"` for points and `"l"` for line. Default is `"l"`. See `plot`
`title`	a string supplying a title for the graphic. If `NULL`, an automatic title is generated
`date.format`	a string indicating the date format for the horizontal axis. Default is `dd/mm/yyyy`. See `strptime` for more options
`new`	if `TRUE` a new graph window is opened
`...`	further options for `plot`

Details

This function and `plot.event` are essentially the same, only some default options differ. There are two of them for convenience only.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Peng, R., Dominici, F. (2008) Statistical methods for environmental epidemiology with R.

See Also

plot,smooth.spline,setup

Examples

```
data(admrio)
setup(admrio,"date")
plot.pollutant(pm10,df=7)
```

`plot.qq` *Plot QQ*

Description

Plot a quantile-quantile graph of the residuals

Usage

```
plot.qq(x, type = "deviance", new = TRUE, ...)
```

Arguments

x	a model fitted by fit.core
type	a quoted string indicating the type of residuals to extract. Default is "deviance". See get.residuals
new	if TRUE a new graph window is opened
...	further options to qqnorm

Details

This function simply interfaces the usual quantile-quantile plot and line of the residuals extracted from x using a single call.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Cook, R. D., Weisberg, S. (1982) *Residuals and Influence in Regression*. Chapman and Hall.

Atkinson, A.C. (1985) *Plots, Transformations and Regression*. Oxford University Press.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

See Also

fit.core,get.residuals

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
plot.qq(m)
```

---

plot.residuals *Plot Residuals*

Description

Extract and/or plot the model residuals

Usage

```
## S3 method for class 'residuals'
plot(x, gaps = FALSE, type = "deviance",
 band=c(-3,3), date.format = "%d/%m/%Y", new = TRUE, ...)
```

Arguments

x	a model fitted by fit.core or a vector of the class residuals
gaps	a logical indicating whether .ares.selection should be honoured
type	a quoted string indicating the type of residuals to extract. Default is "deviance". See get.residuals
band	a vector limits to plot the band. If NULL, the band is omitted. Default is c(-3,3)

date.format	a string indicating the date format for the horizontal axis. Default is dd/mm/yyyy See strptime for more options
new	if TRUE a new graph window is opened
...	further options for plot

Details

If x is a model then the residuals are extracted with get.residuals. If x is a vector of the class residuals the function works as a method for the generic function plot, therefore only the graph is plotted.

Value

A vector of class residuals with extracted and/or adjusted residuals of the model.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>
Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

Pierce, D. A., Schafer, D. W. (1986) Residuals in generalized linear models. *Journal of the American Statistical Association*, **81(396)**,977–986.

See Also

get.residuals,resid,gam,glm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
## extract the residuals and plot
r1 <- plot.residuals(m)

## extract first then plot
r <- get.residuals(m,"adj_deviance")
plot(r)
```

plot.risk *Plot Risk*

Description

Plot estimated risks

Usage

```
## S3 method for class 'risk'
plot(x, labels = rownames(x), new = TRUE, graph.scale = FALSE, ...)
```

Arguments

x	an object of class risk or pdlm.risk output by estimate.risks
labels	a vector of alternate labels for exposures with the same dimension of risks. Default is the vector of rows names of x
new	if TRUE, a new graph window is opened
graph.scale	can be either a logical or a vector with the axis limits. If TRUE or a vector all the graphs will share the same y-axis scale
...	further options for stockplot

Details

This function will plot relative risks computed by estimate.risks. It is seldom called directly.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

plot,estimate.risks

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
rr <- estimate.risks(m,c("pm10","so2"),digits=3,labels=c("PM10","SO2"),method="singlelag",lag.struc=list(l=0:2,

plot(rr)
```

`print.pdlm` *Print a PDL Model*

Description

Print some information on a PDL model

Usage

```
## S3 method for class 'pdlm'
print(x, digits = getOption("digits"), ...)
```

Arguments

`x`	a model of class `pdlm` created by `pdlm`
`digits`	an integer indicating the number of decimal places to print. Default is given by the system option `digits`
`...`	further options for `print`

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>
Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

`pdlm`

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
dlm <- pdlm(m,pm10,lags=5,degrees=2)
print(dlm)
```

`print.risk` *Print Risk*

Description

Print a risk object prettily formatted

Usage

```
## S3 method for class 'risk'
print(x, digits = getOption("digits"), ...)
```

Arguments

`x`	an object of class `risk` or `pdlm.risk` output by `estimate.risks`
`digits`	an integer indicating the number of decimal places to print. Default is given by the system option `digits`
`...`	further options to pass to `print`

Details

This function prints out the risk tables nicely formatted.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

print,estimate.risks

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
rr <- estimate.risks(m,c("pm10","so2"),digits=3,labels=c("PM10","SO2"),method="singlelag",lag.struc=list(l=0:2,

print(rr)
```

---

print.summary *Print Summary*

Description

Print model information

Usage

```
print.summary(x, digits = getOption("digits"), ...)
```

Arguments

x	a model fitted by fit.core
digits	an integer indicating the decimal places. If not supplied, it is taken from options
...	further options for summary.glm or summary.gam

Details

This function prints out both summary.glm and summary.glm outputs added of extra relevant information.

Value

The function invisibly returns a list with the following objects

summary.glm	summary information of the linear part of the model. See summary.glm
summary.gam	summary information of the non-linear part of the model. See summary.gam
dispersion	estimate of the dispersion parameter
pearson	estimate of the generalized Pearson's statistic
residuals.df	residuals degrees of freedom
deviance	estimate of the deviance

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

gam,glm,summary.gam,summary.glm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
print.summary(m)
```

---

print.summary.pdlm *Print Summary of PDL Models*

Description

Method to print a PDLM summary

Usage

```
## S3 method for class 'summary.pdlm'
print(x, digits = getOption("digits"), ...)
```

Arguments

x	an object of class summary.pdlm
digits	an integer indicating the number of decimal places to print. Default is given by the system option digits
...	further option passed by summary.pdlm

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

pdlm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
dlm <- pdlm(m,pm10,lags=5,degrees=2)
summary(dlm)
```

---

resdf *Residual Degrees of Freedom*

Description

Extract residual degrees of freedom from a GAM model

Usage

```
resdf(model)
```

Arguments

model	a model fitted by fit.core

Details

This function extracts necessary information from model. It include both linear and smooth terms.

Value

A scalar with the residual degrees of freedom.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

fit.core,gam,glm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
resdf(m)
```

---

rr.eval *Relative Risk Evaluation*

Description

Compute relative risk for some unit and confidence intervals

Usage

```
rr.eval(beta, se, unit = 1, confidence.level = 0.95)
```

Arguments

beta	regression coefficient of the exposure variate
se	regression standard error of the exposure variate
unit	a value by which the coefficient is multiplied
confidence.level	confidence level for interval computation

Details

The function simply compute the relative risk and the confidence interval for a given amount of pollutant in unit. It is intended to be called internally by estimate.risks.

Value

A vector containing the relative risk and confidence limits.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

McGullagh, P., Nelder, J. A. (1989) *Generalized linear models*. Chapman and Hall.

Hastie, T., Tibshirani, R. (1990) *Generalized additive models*. 2 ed. Chapman and Hall.

See Also

fit.core,plot.risk

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
rr <- estimate.risks(m,c("pm10","so2"),digits=3,labels=c("PM10","SO2"),method="singlelag",lag.struc=list(l=0:2,
```

---

save.plot *Save Plot*

Description

Conveniently save the active plot window to some known graph file format

Usage

```
save.plot(file, width = 520, height = 480)
```

Arguments

file	quoted string with the file name
width	width in pixels
height	height in pixels

Details

This function tries to guess the filter to use from the file name extension. Supported graph formats are Windows meta-file (wmf), portable bitmap (png), jpeg (jpg), portable document format (pdf), encapsulated post script (eps), and scalable vector graphics (svg). If `off` is set to `TRUE`, the active window device is closed after the graph is saved.

Alternatively, one can use the quite new R native `savePlot` that essentially the same job.

This function has been rewritten and GhostScript is not necessary any more.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

`dev.copy`,`dev.copy2eps`

Examples

```
data(admrio)
setup(admrio,"date")
plot.event(resp5)
save.plot("plot_resp5.png")
```

---

setup *Set up ares Environment*

Description

Set up ares environmental by cacching the data, formatting dates, creating holidays, and subsetting on data.

Usage

```
setup(dataset, date.var, selection = NULL, date.format = "%d/%m/%Y",
weekday.ref = "Sun", holidays = TRUE, ...)
```

Arguments

dataset	the data frame for the analysis
date.var	string with the name of the column that holds the date variable. It should be a text variable or a factor
selection	indicator variable for sub-setting. It can be either a quoted string with the name of a column in dataset or a vector
date.format	a string indicating the date format for the horizontal axis. Default is dd/mm/yyyy See strptime for options
weekday.ref	a 3-letter string that match the reference weekday in the U.S. English locale. For example, "Mon" for Monday. Default is "Sun"
holidays	logical indicating if holidays should be created. Default is TRUE. See Details
...	further options for gen.holidays

Details

This function is intended to cache the dataset and set up some environment objects in R so the rest of the library can work properly. Although most of the functions can work stand alone, some of them depend on the object .ares.active.dataset and .ares.selection.

First, .ares.selection is assigned to the global environment, then a date formatted variable is created from date.var and a time index is generated. Factors for days of the week, months, quarters and years are defined. Indicator variables for holidays are created if holidays is set to TRUE. The holidays are created by gen.holidays relying on the date variable. The date variable is defined as doe that stands for "date of event". Country-specific holidays can be be generated on set up time for some countries. One can add these country-specific holidays by adding a file named "XX.hol" to the "/R/library/ares/etc" directory, where "XX" is a 2-letter abbreviation of the country name, for instance "BR.hol" stands for the default option Brazil. The holidays file is self documented. Or you can simply send us an email with the holidays and the dates and we add it for you.

Finally, .ares.active.dataset is assigned to the global environment, every instance of this data frame in the search path is removed and it is reattached.

Attention! The .ares.active.dataset is replaced without warnings. Any changes made to it during a session that was not saved will be lost.

Value

This function invisibly returns .ares.active.dataset as a data frame.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

See Also

attach,search,as.Date

Examples

```
data(admrio)
names(admrio)
setup(admrio,"date")
setup(admrio,"date",selection="warmseason")
```

---

sincos *Sinusoidal Basis Generating Function*

Description

Generate a basis for a sinusoidal function at some fixed period

Usage

```
sincos(period, n = length(.ares.selection), largest.period = 365)
```

Arguments

period	a numeric value indicating the seasonal pattern period
n	an integer indicating the number of observations
largest.period	a numeric value indicating the largest seasonal pattern period. Default is an year or 365 days

Details

This function returns a two-column matrix containing both sine and cosine component evaluated at the given period. The columns are named accordingly.

The function is designed to be used directly from within a formula. Therefore, there is no need to create the sine and cosine vectors prior to the fitting process.

When using with the **ares** library, n may be omitted and it will be taken from the length of .ares.selection.

Value

A two-column matrix containing sine and cosine

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J., Spix, C., Touloumi, G. et al. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol. Community Health* **50 (suppl 1)**, S12–S18.

Box, G., Jenkins, G., Reinsel, G. (1994) *Time Series Analysis : Forecasting and Control*. 3 ed., Prentice Hall.

Diggle, P. J. (1990) *Time Series : A Biostatistical Introduction*. Oxford University Press.

See Also

fit.core,get.residuals

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)+sincos(90)+sincos(180)
m <- fit.core(f)
print.summary(m)
```

---

stockplot *Stock Plot*

Description

Produce a stock type plot

Usage

```
stockplot(mid, low, high, ref.line = NULL, xlabels = seq(1:length(mid)), ticks = 20,
mid.pch = 19, lim.pch = 15, graph.scale = NULL, ...)
```

Arguments

mid	a vector of mid points for the bars, usually mean values
low	a vector of low points for the bars, usually minimum values
high	a vector of high points for the bars, usually maximum values
ref.line	the position on the y-axis to draw a horizontal line. If is set to NULL no line is drawn. Default is NULL
xlabels	a vector of alternate labels for the bars
ticks	an integer indicating the number of tick marks. Default is 20
mid.pch	a number or character indicating the symbol to be used for mid points. Default is a filled circle
lim.pch	a number or character indicating the symbol to be used for end points. Default is a filled square

graph.scale	can be either a logical or a vector with the axis limits. If TRUE or a vector all the graphs will share the same y-axis scale
...	further options for plot

Details

This function implements a stock type plot. Often used to represent stock prices closure, it can indicate the average value at the centre of a bar and minimum and maximum values at the ends of the bar. This plot is very useful to represent statistics and their confidence intervals in a convenient way.

This function is used by plot.risk.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

plot

Examples

```
data(admrio)
setup(admrio,"date")
m <- c(mean(resp5),mean(resp65))
l <- c(min(resp5),min(resp65))
h <- c(max(resp5),max(resp65))
stockplot(m,l,h,xlabels=c("Resp 5","Resp 65"),xlab="Outcome",ylab="Counts")
```

---

summary.pdlm *Summary of PDL Models*

Description

Output the summary of a distributed lag model

Usage

```
## S3 method for class 'pdlm'
summary(object, ...)
```

Arguments

object	a model of class pdlm created by pdlm
...	further options for print.summary.pdlm

Details

The function prints information about the model.

Value

Invisibly returns a list containing

call	function call
coef.table	a table containing coefficients, standard errors, t values and p-values
resid.stats	some statistics on the deviance residuals of the constrained model
deviance	the residual deviance of the constrained model
null.deviance	the null deviance of the constrained model
aic	Akaike Information Criterion of the constrained model
iter	number of Fisher scoring iterations of the constrained model
df.null	null degrees of freedom of the constrained model
df residual	residual degrees of freedom of the constrained model
na	number of observations lost due missing data

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology* **11(3)**, 320–326.

See Also

pdlm

Examples

```
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
dlm <- pdlm(m,pm10,lags=5,degrees=2)
summary(dlm)
```

unload *Unload*

Description

Unload the library and do some housekeeping

Usage

```
unload()
```

Details

If the library is loaded, this function will detach every instance of `.ares.active.dataset` from the search path and remove it from `.GlobalEnv`. Then, it will do some memory cleaning and detach every instance of the library.

After `unload` is called the library has to be reloaded with `library` .

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

See Also

`plot`,`smooth.spline`,`setup`

Examples

```
library(ares)
unload()
```

`whitenoise.test` *White Noise Test*

Description

Run a series of white noise tests

Usage

```
whitenoise.test(object, type = "deviance", k = 25, ...)
```

Arguments

`object`	a model fitted by `fit.core` or a vector of residuals
`type`	a quoted string indicating the type of residuals to extract. Default is `"deviance"`. See `get.residuals`
`k`	an integer indicating the number of lags to compute the partial autocorrelation coefficients
`...`	further options for `ljungbox.test`

Details

This function submits the extracted residuals to a series of white noise tests.

For normality evaluation the robust tests Bownman-Shenton and Jarque-Bera tests are implemented. Besides, Kolmogorov-Smirnov and Shapiro tests are performed using `ks.test` and `shapiro.test` respectively.

Serial dependence of the residuals is evaluated using `ljungbox.test`.

Heteroscedasticity tests will be implemented soon.

Results are printed out on the console.

Value

This function does not return a value.

Author(s)

Washington Junger <wjunger@ims.uerj.br> and Antonio Ponce de Leon <ponce@ims.uerj.br>

Ares-Rio Program at http://www.ims.uerj.br/ares-rio

References

Bowman, K.O. and Shenton L.R. (1975) Omnibus test contours for departures from normality based on b1 and b2. *Biometrika* **62**, 243–250.

Jarque, C. M. and Bera, A. K. (1987) A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, **55(2)**, 163–172.

Spanos, A. (1999) *Probability Theory and Statistical Inference*. Cambridge.

See Also

get.residuals,acf

Examples

```
library(ares)
data(admrio)
setup(admrio,"date")
f <- resp5~s(time)+weekdays+s(tmpmax)+s(humid)
m <- fit.core(f)
r <- get.residuals(m)
whitenoise.test(r)
```

Index