

Package ‘AFR’

March 21, 2022

Type Package

Title Toolkit for Regression Analysis of Kazakhstan Banking Sector Data

Version 0.1.0

Copyright The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market (ARDFM)

Author Timur Abilkassymov [aut],
Shyngys Shuneyev [aut],
Alua Makhmetova [aut, cre]

Maintainer Alua Makhmetova <alua.makhmetova@gmail.com>

Description

Tool is created for regression, prediction and forecast analysis of macroeconomic and credit data. The package includes functions from existing R packages adapted for banking sector of Kazakhstan.

The purpose of the package is to optimize statistical functions for easier interpretation for bank analysts and non-statisticians.

License GPL-2

Depends R (>= 3.5.0)

Imports car, forecast, zoo, regclass, olsrr, stats, lmtest, graphics,
tseries, nlme, ggplot2, gridExtra, utils, rlang, xts, writexl,
mFilter

Suggests knitr, rmarkdown, roxygen2

Repository CRAN

Encoding UTF-8

VignetteBuilder knitr

LazyData true

RoxygenNote 7.1.2

NeedsCompilation no

Date/Publication 2022-03-21 19:00:05 UTC

R topics documented:

ACF	2
adf	3
allpos_helper	4
bg	4
bp	5
checkdata	6
check_betas	6
coeff_length	7
corsel	7
dec_plot	8
difflog	8
gq	9
HP	10
macroKZ	10
opt_size	11
PACF	11
pct1	12
pct4	12
regsel_f	13
reg_plot	14
reg_test	15
vif_reg	15
Index	17

ACF

*Autocorrelation function***Description**

Autocorrelation function demonstrates correlation between stationary time series and its lagged values

Usage

```
ACF(
  x,
  lag.max = NULL,
  type = c("correlation", "covariance", "partial"),
  plot = TRUE,
  na.action = na.contiguous,
  demean = TRUE,
  ...
)
```

Arguments

x	time-series vector
lag.max	maximum lag at which to calculate the ACF
type	character string giving the type of ACF to be computed
plot	logical
na.action	function to be called to handle missing values. na.pass can be used
demean	logical
...	other arguments

Value

An object returns a graph with the x-axis of lags and y-axis of autocorrelation functions. A text gives a hint how to read a graph.

References

Gilbert, P., Plummer, M., Ripley, B.D., Auto- and Cross- Covariance and -Correlation Function Estimation Dancho, Matt. Published 2021-01-18. timetk package

adf	<i>Augmented Dickey Fuller Test</i>
-----	-------------------------------------

Description

ADF test are used to test stationarity of a time-series data

Usage

```
adf(x, k = trunc((length(x) - 1)^(1/3)))
```

Arguments

x	time-series vector
k	the lag order to calculate the test statistic.

Value

A list of values:

p.value	p=value of the test
result	textual result of whether parameter is stationary or not

References

Trapletti, A., Augmented Dickey-Fuller Test Trapletti, A., KPSS Test for Stationarity

allpos_helper	<i>All possible regression internal</i>
---------------	---

Description

Internal function for all possible regression.

Usage

```
allpos_helper(model)
```

Arguments

model	An object of class lm.
-------	------------------------

bg	<i>Breusch-Godfrey test [BG test]</i>
----	---------------------------------------

Description

BG test is used to test for autocorrelation in the errors of a regression model

Usage

```
bg(
  model,
  order = 1,
  order.by = NULL,
  type = c("Chisq", "F"),
  data = list(),
  fill = 0
)
```

Arguments

model	is a (generalized)linear regression model
order	integer. maximal order of serial correlation to be tested.
order.by	Either a vector z or a formula with a single explanatory variable like ~ z
type	the type of test statistic to be returned
data	an optional data frame containing the variables in the model
fill	starting values for the lagged residuals in the auxiliary regression. By default 0 but can also be set to NA.

Value

Textual result of the test. List of values:

LM test	test statistics
p-value	p-value of the BG test

References

Mitchel, D. and Zeileis, A. Published 2021-11-07. lmtest package

bp	<i>Breusch-Pagan test</i>
----	---------------------------

Description

Breusch-Pagan test is used to test against heteroskedasticity of a time-series

Usage

```
bp(model, varformula = NULL, studentize = TRUE, data = list())
```

Arguments

model	is a (generalized)linear regression model
varformula	a formula describing only the potential explanatory variables for the variance (no dependent variable needed). By default the same explanatory variables are taken as in the main regression model.
studentize	logical. If set to TRUE Koenker's studentized version of the test statistic will be used.
data	an optional data frame containing the variables in the model

Value

Textual result of the test. List of values:

BP	test statistics
p-value	p-value of the BP test

References

Torsten, H., Zeileis, A., Farebrother, Richard W., Cummins, C., Millo, G., Mitchell, D., lmtest package Wang, B., 2014, bstats package

checkdata	<i>Data check for errors</i>
-----------	------------------------------

Description

Preliminary check of data frame for missing values, wrong format, outliers.

Usage

```
checkdata(x)
```

Arguments

x is a data frame

Value

A matrix of logical and numerical values for the dataset.

check_betas	<i>All possible regression variable coefficients</i>
-------------	--

Description

Returns the coefficients for each variable from each model.

Usage

```
check_betas(object, ...)
```

Arguments

object An object of class lm.
 ... Other arguments.

Value

check_betas returns a data.frame containing:

model_index	model number
predictor	predictor
beta_coef	coefficient for the predictor

References

Hebbali, Aravind. Published 2020-02-10. olssr package

Examples

```
model <- lm(real_gdp~imp+exp+usdkzt+eurkzt, data = macroKZ)
check_betas(model)
```

coeff_length	<i>Coefficients length</i>
--------------	----------------------------

Description

Returns the length of the coefficient names.

Usage

```
coeff_length(predicts, gap)
```

Arguments

predicts	Name of the predictors in the model.
gap	A numeric vector.

Value

no return value

corse1	<i>Multicollinearity test</i>
--------	-------------------------------

Description

multicollinearity is the occurrence of high interrelations among two or more independent variables in a multiple regression.

Usage

```
corse1(x, thrs, num)
```

Arguments

x	is a numeric vector or matrix
thrs	threshold set to calculate correlation above
num	logical

Value

A matrix of numerical and logical values of correlation between variables.

dec_plot *Decomposition plot*

Description

The function depicts decomposition of regressors as a stacked barplot

Usage

```
dec_plot(model, dataset, print_plot = TRUE)
```

Arguments

model	An object of class lm.
dataset	A dataset based on which model was built
print_plot	logical

Value

A plot that shows decomposition of the model.

Author(s)

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market (AFR)

References

Hebbali, Aravind. Published 2020-02-10. olssr package

Examples

```
model <- lm(real_gdp ~ usdkzt + eurkzt + imp+exp, data = macroKZ)
dec_plot(model, macroKZ)
```

difflog *Transforming time-series data to stationary*

Description

Difference of logarithms is finding the difference between two consecutive logarithm values of a time-series

Usage

```
difflog(x, lag = 1, difference = 1)
```


Arguments

x	time-series vector
lag	lagged period
difference	difference between x items

Value

A length shorter than the original length. A returned value is a numerical.

gq	<i>Godfrey-Quandt test</i>
----	----------------------------

Description

Godfrey-Quandt test is used to test against heteroskedasticity of a time-series

Usage

```
gq(
  model,
  point = 0.5,
  fraction = 0,
  alternative = c("greater", "two.sided", "less"),
  order.by = NULL,
  data = list()
)
```

Arguments

model	is a (generalized)linear regression model
point	numerical. If point is smaller than 1 it is interpreted as percentages of data
fraction	numerical. The number of central observations to be omitted.
alternative	a character string specifying the alternative hypothesis.
order.by	Either a vector z or a formula with a single explanatory variable like ~ z
data	an optional data frame containing the variables in the model.

Value

Textual result of the test. List of values:

GQ	test statistics
p-value	p-value of the BP test

References

Torsten, H., Zeileis, A., Farebrother, Richard W., Cummins, C., Millo, G., Mitchell, D., lmtest package Wang, B., 2014, bstats package

HP *Hodrick-Prescott filter*

Description

Hodrick-Prescott filter is a data smoothing technique that removes trending in time series data frame

Usage

```
HP(x, freq = NULL, type = c("lambda", "frequency"), drift = FALSE)
```

Arguments

x	time-series vector
freq	integer
type	character, indicating the filter type
drift	logical

Value

A "mFilter" object of the mFilter package.

macroKZ *macroKZ dataset*

Description

A time series data frame of 48 quarterly observations of 50 macroeconomic and 10 financial parameters for 2010-2021 period.

Usage

```
macroKZ
```

Source

Bureau of National statistics, Agency for Strategic planning and reforms of the Republic of Kazakhstan

References

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market

opt_size	<i>Necessary size of the time-series dataset</i>
----------	--

Description

Estimates number of models generated from given number of regressors x

Usage

```
opt_size(model)
```

Arguments

model is a linear regression model a class lm.

Value

A numerical value of the number of models.

PACF	<i>Partial autocorrelation function</i>
------	---

Description

(partial) Autocorrelation function demonstrates (partial) correlation between stationary time series and its lagged values

Usage

```
PACF(
  x,
  lag.max = NULL,
  plot = TRUE,
  na.action = na.contiguous,
  demean = TRUE,
  ...
)
```

Arguments

x	time-series vector
lag.max	maximum lag at which to calculate the ACF
plot	logical
na.action	function to be called to handle missing values. na.pass can be used
demean	logical
...	other arguments

Value

An object returns a graph with the x-axis of lags and y-axis of autocorrelation functions. A text gives a hint how to read a graph.

References

Gilbert, P., Plummer, M., Ripley, B.D., Auto- and Cross- Covariance and -Correlation Function Estimation Dancho, Matt. Published 2021-01-18. timetk package

pct1 *Transforming time-series data to stationary*

Description

Percent change is a change between two consecutive terms,

Usage

pct1(x)

Arguments

x time-series vector(s)

Value

A vector of length shorter than the original vector. Returned value is a numeric.

pct4 *Transforming time-series data to stationary*

Description

Percent change is a change between a term and its lagged value for prior period,

Usage

pct4(x)

Arguments

x time-series vector(s)

Value

A vector of length shorter than the original vector. Returned value is a numeric.

regsel_f	<i>Regressors selection</i>
----------	-----------------------------

Description

The function allows to choose regressors based on multiple criteria as AIC, RMSE etc

Usage

```
regsel_f(
  model,
  pval = 0.3,
  metric = "adjr" & "aic",
  progress = FALSE,
  details = FALSE,
  ...
)

## S3 method for class 'regsel_f'
print(x, ...)

## S3 method for class 'regsel_f'
plot(x, print_plot = TRUE, ...)
```

Arguments

model	is a linear regression model
pval	p value; variables with p value less than pval will enter into the model
metric	statistical metrics used to estimate the best model
progress	Logical; if TRUE, will display variable selection progress.
details	Logical; if TRUE, will print the regression result at each step.
...	other arguments
x	An object.
print_plot	logical; if TRUE, prints the plot else returns a plot object.

Value

An object containing the following components:

model	final model; an object of class <code>lm</code>
steps	number of steps
predictors	variables added to the model
rsquare	coefficient of determination
aic	akaike information criteria

sbc	bayesian information criteria
sbic	sawa's bayesian information criteria
adjr	adjusted r-square
rmse	root mean square error
mallows_cp	mallow's Cp
indvar	predictors

References

Hebbali, Aravind. Published 2020-02-10. olssr package

reg_plot	<i>Regression forecast plot</i>
----------	---------------------------------

Description

The function depicts forecast and actual data.

Usage

```
reg_plot(model, dataset)
```

Arguments

model	An object of class lm.
dataset	A dataset based on which model was built.

Value

A plot that shows predictive and actual values.

Author(s)

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market (AFR)

Examples

```
model <- lm(real_gdp ~ usdkzt + eurkzt + imp+exp, data = macroKZ)
reg_plot(model, macroKZ)
```

reg_test	<i>Test for detecting violation of Gauss-Markov assumptions.</i>
----------	--

Description

Test for detecting violation of Gauss-Markov assumptions.

Usage

```
reg_test(y)
```

Arguments

y A numeric vector or an object of class lm.

Value

reg_test returns an object of class "reg_test". An object of class "reg_test" is a list containing the following components:

bp	Breusch-Pagan statistic
bg	Breusch-Godfrey statistic
dw	Durbin-Watson statistic
gq	Godfrey-Quandt statistic

Examples

```
model <- lm(real_gdp~imp+exp+poil+eurkzt+usdkzt, macroKZ)
reg_test(model)
```

vif_reg	<i>VIF by variable</i>
---------	------------------------

Description

Calculates the variation inflation factors of all predictors in regression models

Usage

```
vif_reg(model)
```

Arguments

model is a linear regression model

Value

A data frame with columns for variable and VIF result. Textual result for VIF test presents whether variables pass the test or not.

References

Petrie, Adam. Published 2020-02-21. regclass package

Index

* datasets

macroKZ, [10](#)

ACF, [2](#)

adf, [3](#)

allpos_helper, [4](#)

bg, [4](#)

bp, [5](#)

check_betas, [6](#)

checkdata, [6](#)

coeff_length, [7](#)

corsel, [7](#)

dec_plot, [8](#)

difflog, [8](#)

gq, [9](#)

HP, [10](#)

macroKZ, [10](#)

opt_size, [11](#)

PACF, [11](#)

pct1, [12](#)

pct4, [12](#)

plot.regsel_f (regsel_f), [13](#)

print.regsel_f (regsel_f), [13](#)

reg_plot, [14](#)

reg_test, [15](#)

regsel_f, [13](#)

vif_reg, [15](#)