Package ‘AFR’

October 29, 2023

Type Package
Title Toolkit for Regression Analysis of Kazakhstan Banking Sector Data
Version 0.3.5
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, olsrr, stats, lmtest, graphics, nlme, ggplot2, tseries, gridExtra, utils, rlang, xts, nortest, goftest, cli
Suggests knitr, rmarkdown, roxygen2
SystemRequirements C++
Repository CRAN
Encoding UTF-8
VignetteBuilder knitr
LazyData true
RoxygenNote 7.1.2
NeedsCompilation no
Date/Publication 2023-10-29 13:40:02 UTC
R topics documented:

bg ......................................................... 2
bp ......................................................... 3
checkdata ........................................... 4
check_betas .......................................... 4
corSel ................................................ 5
dec_plot ............................................. 5
difflog ............................................... 6
finratKZ ............................................. 7
gq ....................................................... 8
HP ....................................................... 9
macroKZ ............................................ 9
ols_test_normality ................................ 12
opt_size .......................................... 12
pct1 .................................................. 13
pct4 .................................................. 13
pt_multi .......................................... 14
pt_one ............................................. 14
regsel_f .......................................... 15
reg_plot .......................................... 16
reg_test .......................................... 16
vif_reg ........................................... 17

Index 18

bg  Breusch-Godfrey test [BG test]

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
)
**bp**

**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 \( \sim z \)
- `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.

**References**

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

**Examples**

```r
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
bp(model)
```

---

**Description**

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

**Usage**

```r
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

**References**


**Examples**

```r
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
bp(model)
```
**Description**

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

Missing items: The number of missing values in each column of the dataset. Numeric format: The number of non-numeric variables in each column of the dataset. Outliers: The number of outliers in each column of the dataset.

**Usage**

```r
checkdata(x)
```

**Arguments**

- `x` is a data frame

**Examples**

```r
data(macroKZ)
checkdata(macroKZ)
```

---

**Description**

Returns the coefficients for each variable from each model.

**Usage**

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

**Arguments**

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

**Value**

`check_betas` returns a data.frame containing:

- `x` model
References

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

Examples

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

corsel

**Multicollinearity test**

description

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

Usage

corsel(x, thrs, num)

Arguments

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

Examples

data(macroKZ)
corsel(macroKZ,num=FALSE,thrs=0.65)

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 \texttt{lm}.
dataset A dataset based on which model was built
print_plot logical

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
\begin{verbatim}
data(macroKZ)
model <- lm(real_gdp ~ usdkzt + eurkzt + imp+exp, data = macroKZ)
dec_plot(model, macroKZ)
\end{verbatim}

difflog \hspace{1cm} 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

Examples
\begin{verbatim}
data (macroKZ)
new<-pct1(macroKZ)
\end{verbatim}
Description

finratKZ dataset

Usage

finratKZ

Format

Dataset of 400 corporate borrowers, i.e. 200 standard (IFRS stage 1) and 200 default ones, characterized by 29 financial ratios.

**Default**  Dummy variable where 0 - standard (IFRS stage 1) borrower, 1 - default borrower
**Rev_gr**  Revenue growth rate
**EBITDA_gr**  EBITDA growth rate
**Cap_gr**  Capital growth rate
**CR**  Current ratio
**QR**  Quick ratio
**Cash_ratio**  Cash ratio
**WC_cycle**  Working capital cycle
**DTA**  Debt-to-assets
**DTE**  Debt-to-equity
**LR**  Leverage ratio (Total assets/Total equity)
**EBITDA_debt**  EBITDA-to-debt
**IC**  Interest coverage (Income statement)
**CTI**  Cash-to-income
**IC_CF**  Interest coverage (Cash flow statement)
**DCR**  Debt coverage ratio (Cash flow from operations/Total debt)
**CFR**  Cash flow to revenue
**CRA**  Cash return on assets (Cash flow from operations/Total assets)
**CRE**  Cash return on equity (Cash flow from operations/Total equity)
**ROA**  Return on assets
**ROE**  Return on equity
**NPM**  Net profit margin
**GPM**  Gross profit margin
**OPM**  Operating profit margin
RecT  Receivables turnover
InvT  Inventory turnover
PayT  Payables turnover
TA    Total assets turnover
FA    Fixed assets turnover
WC    Working capital turnover

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

---

gq

**Godfrey-Quandt test**

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.

References

Examples
```
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
gq(model)
```
Hodrick-Prescott filter for time series data

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

Examples

data(macroKZ)
HP(macroKZ[,2])

---

macroKZ dataset

Description

macroKZ dataset

Usage

macroKZ

Format

A time series data frame of 54 quarterly observations of 50 macroeconomic and 10 financial parameters for 2010-2023 period.

- real_gdp: Real GDP
- GDD_Agr_R: Real gross value added Agriculture
- GDD_Min_R: Real gross value added Mining
- GDD_Man_R: Real gross value added Manufacture
GDD_Elc_R  Real gross value added Electricity
GDD_Con_R  Real gross value added Construction
GDD_Trd_R  Real gross value added Trade
GDD_Trn_R  Real gross value added Transportation
GDD_Inf_R  Real gross value added Information
GDD_R  Real gross value added
GDP_DEF  GDP deflator
Rincpop_q  Real population average monthly income
Rexpop_q  Real population average monthly expenses
Rwage_q  Real population average monthly wage
imp  Import
exp  Export
cpi  Inflation
realest_resed_prim  Real price for estate in primary market
realest_resed_sec  Real price for estate in secondary market
realest_comm  Real price for commercial estate
index_stock_weighted  Change in stock value for traded companies
ntrade_Agr  Change in stock value for non-traded companies Agriculture
ntrade_Min  Change in stock value for non-traded companies Mining
ntrade_Man  Change in stock value for non-traded companies Manufacture
ntrade_Elc  Change in stock value for non-traded companies Electricity
ntrade_Con  Change in stock value for non-traded companies Construction
ntrade_Trd  Change in stock value for non-traded companies Trade
ntrade_Trn  Change in stock value for non-traded companies Transportation
ntrade_Inf  Change in stock value for non-traded companies Information
fed_fund_rate  Federal Funds Rate
govsec_rate_kzt_3m  Return on government securities in KZT, 3 m
govsec_rate_kzt_1y  Return on government securities in KZT, 1 year
govsec_rate_kzt_7y  Return on government securities in KZT, 7 years
govsec_rate_kzt_10y  Return on government securities in KZT, 10 years
tonia_rate  TONIA
rate_kzt_mort_0y_1y  Weighted average mortgage lending rate for new loans, less than a year
rate_kzt_mort_1y_iy  Weighted average mortgage lending rate for new loans, more than a year
rate_kzt_corp_0y_1y  Weighted average mortgage lending rate for new loans to non-financial organizations in KZT, less than a year
rate_usd_corp_0y_1y  Weighted average mortgage lending rate for new loans to non-financial organizations in CKB, less than a year
rate_kzt_corp_1y_iy  Weighted average mortgage lending rate for new loans to non-financial organizations in KZT, more than a year
rate_usd_corp_1y_iy  Weighted average mortgage lending rate for new loans to non-financial organizations in CKB, more than a year
rate_kzt_indv_0y_1y  Weighted average mortgage lending rate for consumer loans in KZT, less than a year
rate_kzt_indv_1y_iy  Weighted average mortgage lending rate for consumer loans in KZT, less than a year
usdkzt  USD KZT exchange rate
eurkzt  EUR KZT exchange rate
rurkzt  RUB KZT exchange rate
poil  Price for Brent
realest_resed_prim_rus  Real price for estate in primary market in Russia
realest_resed_sec_rus  Real price for estate in secondary market in Russia
cred_portfolio  credit portfolio
coef_liq_k4  k4 prudential coefficient
coef_k1  k1 prudential coefficient
coef_k3  k3 prudential coefficient
provisions  provisions
percent_margin  percent margin
com_inc  commissionary income
com_exp  commissionary expenses
oper_inc  operational income
oth_inc  other income
DR  default rate

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
ols_test_normality

Test for normality Test for detecting violation of normality assumption.

Description

Test for normality Test for detecting violation of normality assumption.

Usage

ols_test_normality(model, ...)

Arguments

model an object of class lm.
... Other arguments.

Value

ols_test_normality is a list containing the following components:

kolmogorv kolmogorov smirnov statistic
shapiro shapiro wilk statistic
cramer cramer von mises statistic
anderson anderson darling statistic

Examples

data(macroKZ)
model <- lm(real_gdp ~ imp + exp + usdkzt + poil, data = macroKZ)
ols_test_normality(model)

opt_size

Necessary size of the time-series dataset

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.
Examples

```r
data(macroKZ)
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
opt_size(model)
```

---

\textit{pct1}  
\textit{Transforming time-series data to stationary}

Description

Percent change is a change between two consecutive terms.

Usage

\texttt{pct1(x)}

Arguments

\texttt{x} \hspace{2cm} \text{time-series vector(s)}

Examples

```r
data (macroKZ)
new<-pct1(macroKZ)
```

---

\textit{pct4}  
\textit{Transforming time-series data to stationary}

Description

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

Usage

\texttt{pct4(x)}

Arguments

\texttt{x} \hspace{2cm} \text{time-series vector(s)}

Examples

```r
data (macroKZ)
new<-pct4(macroKZ)
```
pt_multi

Pluto-Tasche method for multi-year probability of default (PD) analysis

Description
Calculates the variation inflation factors of all predictors in regression models

Usage
pt_multi(pf, num_def, conf_level, num_years)

Arguments
- **pf**: unconditional portfolio distribution from the worst to the best credit quality
- **num_def**: number of defaults in a given rating class
- **conf_level**: confidence interval of PD estimates
- **num_years**: number of periods used in the PD estimation

Examples
```r
pf <- c(10,20,30,40)
num_def <- c(1,2,3,4)
conf_level = 0.99
num_years = 3
pt_multi(pf, num_def, conf_level, num_years)
```

pt_one

Pluto-Tasche method for one-year probability of default (PD) analysis

Description
Calculates probability of default according to One-period Pluto and Tasche model

Usage
pt_one(pf, num_def, ci = 0.9)

Arguments
- **pf**: unconditional portfolio distribution from the worst to the best credit quality
- **num_def**: number of defaults in a given rating class
- **ci**: confidence interval of PD estimates
regsel_f

References

Examples
pf <- c(10, 20, 30, 40)
num_def <- c(1, 2, 3, 4)
pt_one(pf, num_def, ci= 0.9)

__________________________

regsel_f Regressors selection
__________________________

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,
  ...
)

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.
...

References
Hebbali, Aravind. Published 2020-02-10. olsr package

Examples
data(macroKZ)
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
regsel_f(model)
**Reg_plot**  
*Regression forecast plot*

**Description**

The function depicts forecast and actual data.

**Usage**

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

**Arguments**

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

**Author(s)**

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

**Examples**

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

---

**Reg_test**  
*Test for detecting violation of Gauss-Markov assumptions.*

**Description**

Test for detecting violation of Gauss-Markov assumptions.

**Usage**

```r
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**

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

---

**Description**

Calculates the variation inflation factors of all predictors in regression models.

**Usage**

vif_reg(model)

**Arguments**

- **model**: is a linear regression model

**References**

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

**Examples**

data(macroKZ)
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
vif_reg(model)
Index

* datasets
  finratKZ, 7
  macroKZ, 9

bg, 2
bp, 3

check_betas, 4
checkdata, 4
corsel, 5
dec_plot, 5
difflog, 6

finratKZ, 7
gq, 8
HP, 9

macroKZ, 9

ols_test_normality, 12
opt_size, 12

pct1, 13
pct4, 13
pt_multi, 14
pt_one, 14

reg_plot, 16
reg_test, 16
regsel_f, 15

vif_reg, 17