Package 'lindia'

October 13, 2022

Type Package

Title Automated Linear Regression Diagnostic

Version 0.9

Description Provides a set of streamlined functions that allow easy generation of linear regression diagnostic plots necessarily for checking linear model assumptions.
This package is meant for easy scheming of linear regression diagnostics, while preserving merits of ``The Grammar of Graphics'' as implemented in 'ggplot2'.
See the 'ggplot2' website for more information regarding the specific capability of graphics.

Depends R (>= 3.2.2)

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Imports MASS, ggplot2, gridExtra

URL https://github.com/yeukyul/lindia

BugReports https://github.com/yeukyul/lindia/issues

NeedsCompilation no

Author Yeuk Yu Lee [aut, cre],
Samuel Ventura [aut]

Maintainer Yeuk Yu Lee <lyeukyu@gmail.com>

Repository CRAN

Date/Publication 2017-05-10 02:29:20 UTC

R topics documented:

---

gg_boxcox *Plot boxcox graph in ggplot with suggested lambda transformation*

Description

Plot boxcox graph in ggplot with suggested lambda transformation

Usage

```
gg_boxcox(fitted.lm, showlambda = TRUE, lambdaSF = 3, scale.factor = 0.5)
```

Arguments

`fitted.lm`	a fitted linear model (i.e. lm, glm) that contains fitted regression
`showlambda`	logical; controls whether lambda value should be displayed on graph. Defaults to TRUE
`lambdaSF`	numeric; controls to how many significant figure is lambda rounded to. Defaults to 3.
`scale.factor`	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 0.5.

Value

A ggplot object that contains boxcox graph

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price ~ Passengers + Length + RPM, data = Cars93)
gg_boxcox(cars_lm)
```

gg_cooksd *Plot cook's distance graph*

Description

Plot cook's distance graph

Usage

```
gg_cooksd(fitted.lm, label = TRUE, show.threshold = TRUE,
  threshold = "convention", scale.factor = 0.5)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
label	logical; whether or not to label observation number larger than threshold. Default to TRUE.
show.threshold	logical; determine whether or not threshold line is to be shown. Default to TRUE.
threshold	string; determining the cut off label of cook's distance. Choices are "baseR" (0.5 and 1), "matlab" (mean(cooksd)*3), and "convention" (4/n and 1). Default to "convention".
scale.factor	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 0.5.

Value

A ggplot object that contains a cook's distance plot

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price ~ Passengers + Length + RPM, data = Cars93)
gg_cooksd(cars_lm)
```

gg_diagnose *Plot all diagnostic plots given fitted linear regression line.*

Description

Plot all diagnostic plots given fitted linear regression line.

Usage

```
gg_diagnose(fitted.lm, theme = NULL, ncol = NA, plot.all = TRUE,
  scale.factor = 0.5, boxcox = FALSE, max.per.page = NA)
```

Arguments

fitted.lm	lm object that contains fitted regression
theme	ggplot graphing style using 'ggplot::theme()'. A ggplot graphing style to apply to all plots. Default to null.
ncol	specify number of columns in resulting plot per page. Default to make a square matrix of the output.
plot.all	logical; determine whether plot will be returned as an arranged grid. When set to false, the function will return a list of diagnostic plots. Parameter defaults to TRUE.
scale.factor	numeric; scales the point size, linewidth, labels in all diagnostic plots to allow optimal viewing. Defaults to 0.5.
boxcox	logical; detemine whether boxcox plot will be included. Parameter defaults to FALSE.
max.per.page	numeric; maximum number of plots allowed in one page.

Value

An arranged grid of linear model diagnostics plots in ggplot. If plot.all is set to FALSE, a list of ggplot objects will be returned instead. Name of the plots are set to respective variable names.

Examples

```
library(MASS)
data(Cars93)
# a regression with categorical variable
cars_lm <- lm(Price ~ Passengers + Length + RPM + Origin, data = Cars93)
gg_diagnose(cars_lm)
# customize which diagnostic plot is included
plots <- gg_diagnose(cars_lm, plot.all = FALSE)
names(plots)     # get name of the plots
exclude_plots <- plots[-c(1, 3) ]    #exclude certain diagnostics plots
include_plots <- plots[c(1, 3)]      # include certain diagnostics plots
plot_all(exclude_plots)              # make use of plot_all() in lindia
plot_all(include_plots)
```

gg_qqplot *Plot quantile-quantile plot (QQPlot) in ggplot with qqline shown.*

Description

Plot quantile-quantile plot (QQPlot) in ggplot with qqline shown.

Usage

```
gg_qqplot(fitted.lm, scale.factor = 1)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
scale.factor	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 1.

Value

A qqplot with fitted qqline

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price ~ Passengers + Length + RPM, data = Cars93)
gg_qqplot(cars_lm)
```

gg_resfitted *Generate residual plot of residuals against fitted value*

Description

Generate residual plot of residuals against fitted value

Usage

```
gg_resfitted(fitted.lm, scale.factor = 1)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
scale.factor	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 1.

Value

A ggplot object

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price ~ Passengers + Length + RPM, data = Cars93)
gg_resfitted(cars_lm)
```

---

gg_reshist *Generate histogram of residuals in ggplot.*

Description

Generate histogram of residuals in ggplot.

Usage

```
gg_reshist(fitted.lm, bins = NULL)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
bins	bin size for histogram

Value

A ggplot object

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price~ Passengers + Length + RPM, data = Cars93)
gg_reshist(cars_lm)
# specify number of bins
gg_reshist(cars_lm, bins = 20)
```

gg_resleverage *Plot residual versus leverage plot in ggplot.*

Description

Plot residual versus leverage plot in ggplot.

Usage

```
gg_resleverage(fitted.lm, method = "loess", se = FALSE, scale.factor = 1)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
method	smoothing method of fitted line on scale-location plot. eg. "lm", "glm", "gam", "loess", "rlm". See http://docs.ggplot2.org/current/geom_smooth.html for more details.
se	logical; determines whether se belt should be plotted on plot
scale.factor	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 1.

Value

A ggplot object that contains residual vs. leverage graph

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price ~ Passengers + Length + RPM, data = Cars93)
gg_resleverage(cars_lm)
```

gg_resX *Generate residual plot of residuals against predictors*

Description

Generate residual plot of residuals against predictors

Usage

```
gg_resX(fitted.lm, plot.all = TRUE, scale.factor = 0.5, max.per.page = NA,
  ncol = NA)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
plot.all	boolean value to determine whether plot will be return as a plot arranged using 'grid.arrange()'. When set to false, the function would return a list of residual plots. Parameter defaults to TRUE.
scale.factor	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 0.5.
max.per.page	numeric; maximum number of plots allowed in one page. Parameter defaults to fit all plots on one page.
ncol	specify number of columns in resulting plot per page. Default to make a square matrix of the output.

Value

An arranged grid of residuals against predictor values plots in ggplot. If plotall is set to FALSE, a list of ggplot objects will be returned instead. Name of the plots are set to respective variable names.

Examples

```
library(MASS)
data(Cars93)
# a regression with categorical variable
cars_lm <- lm(Price ~ Passengers + Length + RPM + Origin, data = Cars93)
gg_resX(cars_lm)
# customize which diagnostic plot is included by have gg_resX to return a list of plots
plots <- gg_resX(cars_lm, plot.all = FALSE)
names(plots)     # get name of the plots
exclude_plots <- plots[-1 ]    #exclude certain residual plots
include_plots <- plots[1]      # include certain residual plots
plot_all(exclude_plots)       # make use of plot_all() in lindia
plot_all(include_plots)
```

---

gg_scalelocation *Plot scale-location (also called spread-location plot) in ggplot.*

Description

Plot scale-location (also called spread-location plot) in ggplot.

Usage

```
gg_scalelocation(fitted.lm, method = "loess", scale.factor = 1,
  se = FALSE)
```

Arguments

fitted.lm	a fitted linear model (i.e. lm, glm) that contains fitted regression
method	smoothing method of fitted line on scale-location plot. eg. "lm", "glm", "gam", "loess", "rlm". See http://docs.ggplot2.org/current/geom_smooth.html for more details.
scale.factor	numeric; scales the point size and linewidth to allow customized viewing. Defaults to 1.
se	logical; determines whether se belt should be plotted on plot

Value

A ggplot object that contains scale-location graph

Examples

```
library(MASS)
data(Cars93)
cars_lm <- lm(Price ~ Passengers + Length + RPM, data = Cars93)
gg_scalelocation(cars_lm)
```

---

plot_all *Plot all given plots in a square matrix form.*

Description

Plot all given plots in a square matrix form.

Usage

```
plot_all(plots, ncol = NA, max.per.page = NA)
```

Arguments

plots	a list of plots
ncol	numeric; the number of column that the arranged grid need to be. defaults to fitting all plots in square matrix
max.per.page	numeric; maximum number of plots allowed in one page.

Value

plots in a given list arrangeed using gridExtra

Examples

```
library(MASS)
data(Cars93)
# a regression with categorical variable
cars_lm <- lm(Price ~ Passengers + Length + RPM + Origin, data = Cars93)
plots <- gg_diagnose(cars_lm, plot.all = FALSE)
names(plots)
selected.plots <- plots[-c(2, 5)]
plot_all(selected.plots)
```

Index