Package ‘kvr2’

February 12, 2026

Type Package

Title Calculate and Compare Multiple Definitions of Coefficient of Determination

Version 0.1.0

Description Calculate nine types of coefficients of determination (R-squared) based on the classification by Kvalseth (1985) <doi:10.1080/00031305.1985.10479448>.
This package is designed for educational purposes to demonstrate how R-squared values can fluctuate depending on the choice of formula, particularly in power regression models or linear models without an intercept.
By providing a comprehensive list of definitions, it helps users understand the mathematical sensitivity of goodness-of-fit indices.

URL https://github.com/indenkun/kvr2, https://indenkun.github.io/kvr2/

BugReports https://github.com/indenkun/kvr2/issues

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Imports stats

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Mao Kobayashi [aut, cre]

Maintainer Mao Kobayashi <kobamao.jp@gmail.com>

Repository CRAN

Date/Publication 2026-02-12 08:20:25 UTC

Contents

comp_fit 2
print.r2_kvr2 4
r2 4
r2_adjusted 7

Index **9**

`comp_fit` *Calculate Comparative Fit Measures for Regression Models*

Description

Calculates goodness-of-fit metrics based on Kvalseth (1985), including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Squared Error (MSE). This function provides a unified output for comparing different model specifications.

Usage

```
comp_fit(model, type = c("auto", "liner", "power"))

RMSE(model, type = c("auto", "liner", "power"))

MAE(model, type = c("auto", "liner", "power"))

MSE(model, type = c("auto", "liner", "power"))
```

Arguments

`model`	A linear model or power regression model of the `lm`.
`type`	Character string. Selects the model type: `"linear"`, `"power"`, or `"auto"` (default). In `"auto"`, the function detects if the dependent variable is log-transformed.

Details

The metrics are calculated according to the formulas in Kvalseth (1985):

- **RMSE**: Root Mean Squared Residual or Error

$$RMSE = \sqrt{\frac{\sum(y - \hat{y})^2}{n}}$$

- **MAE**: Mean Absolute Residual or Error

$$MAE = \frac{\sum |y - \hat{y}|}{n}$$

- **MSE**: Mean Squared Residual or Error (Adjusted for degrees of freedom)

$$MSE = \frac{\sum(y - \hat{y})^2}{n - p}$$

where n is the sample size and p is the number of model parameters (including the intercept).

Note on MSE: In many modern contexts, "MSE" refers to the mean squared error without degree-of-freedom adjustment (denominator n). However, this function follows Kvalseth's definition, which uses $n - p$ as the denominator.

Value

An object of class `comp_kvr2`, which is a list containing the calculated RMSE, MAE, and MSE values.

Note

The power regression model must be based on a logarithmic transformation.

The auto-selection between linear regression and power regression models is determined by whether the dependent variable's name contains "log". If the name "log" is intentionally used for a linear regression model, the selection cannot be made correctly.

References

Tarald O. Kvalseth (1985) Cautionary Note about R 2 , The American Statistician, 39:4, 279-285, doi: 10.1080/00031305.1985.10479448

See Also

`print.comp_kvr2()`

Examples

```
# example data set 1. Kvålseth (1985).
df1 <- data.frame(x = c(1:6),
                  y = c(15,37,52,59,83,92))
model_intercept <- lm(y ~ x, df1)
model_without <- lm(y ~ x - 1, df1)
model_power <- lm(log(y) ~ log(x), df1)
comp_fit(model_intercept)
comp_fit(model_without)
comp_fit(model_power)
```

print.r2_kvr2 *Print Methods for r2 and comp_fit calculation Objects*

Description

Printing objects of class "r2_kvr2" (generated by r2()) or "comp_kvr2" (generated by comp_fit()), respectively, by simple print methods.

Usage

```
## S3 method for class 'r2_kvr2'
print(x, ..., digits = 4)

## S3 method for class 'comp_kvr2'
print(x, ..., digits = 4)
```

Arguments

x	AN object of class "r2_kvr2" or "comp_kvr2".
...	Further arguments passed to or from other methods.
digits	The number of significant digits to be used for printing. Default to 4.

Details

These methods format the calculated statistics into a human-readable summary, displaying each index or metric with its corresponding value.

Value

The input object is returned invisibly (via invisible(x)). This function is called for its side effect of printing the results of r2() or comp_fit() calculations to the console.

See Also

r2() comp_fit() r2_adjusted()

r2 *Calculate Multiple Definitions of Coefficient of Determination (R-squared)*

Description

Calculates nine types of coefficients of determination (R^2) based on the classification by Kvalseth (1985). This function is designed to demonstrate how R^2 values can vary depending on their mathematical definition, particularly in models without an intercept or in power regression models

Usage

```
r2(model, type = c("auto", "liner", "power"), adjusted = FALSE)

r2_1(model, type = c("auto", "liner", "power"))

r2_2(model, type = c("auto", "liner", "power"))

r2_3(model, type = c("auto", "liner", "power"))

r2_4(model, type = c("auto", "liner", "power"))

r2_5(model, type = c("auto", "liner", "power"))

r2_6(model, type = c("auto", "liner", "power"))

r2_7(model, type = c("auto", "liner", "power"))

r2_8(model, type = c("auto", "liner", "power"))

r2_9(model, type = c("auto", "liner", "power"))
```

Arguments

`model`	A linear model or power regression model of the `lm`.
`type`	Character string. Selects the model type: `"linear"`, `"power"`, or `"auto"` (default). In `"auto"`, the function detects if the dependent variable is log-transformed.
`adjusted`	Logical. If `TRUE`, calculates the adjusted coefficient of determination for each formula.

Details

The nine coefficient equations from R_1^2 to R_9^2 are based on Kvalseth (1985) and are as follows:

- R_1^2: Proportion of total variance explained.

$$R_1^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2}$$

- R_2^2: Based on the variation of predicted values.

$$R_2^2 = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y - \bar{y})^2}$$

- R_3^2: Ratio of variation using the mean of predicted values.

$$R_3^2 = \frac{\sum(\hat{y} - \bar{\hat{y}})^2}{\sum(y - \bar{y})^2}$$

- R_4^2: Incorporates the mean residual.

$$R_4^2 = 1 - \frac{\sum(e - \bar{e})^2}{\sum(y - \bar{y})^2}, \quad e = y - \hat{y}$$

- R_5^2: The square of the multiple correlation coefficient between the dependent variable and the independent variable (a comprehensive indicator in linearized models).

$$R_5^2 = \text{squared multiple correlation coefficient between the regressand and the regressors}$$

- R_6^2: Square of Pearson's correlation coefficient between observed y and predicted $\hat{y}$.

$$R_6^2 = \frac{\left(\sum(y - \bar{y})(\hat{y} - \bar{\hat{y}})\right)^2}{\sum(y - \bar{y})^2 \sum(\hat{y} - \bar{\hat{y}})^2}$$

- R_7^2: Recommended for models without an intercept.

$$R_7^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum y^2}$$

- R_8^2: Alternative form for models without an intercept.

$$R_8^2 = \frac{\sum \hat{y}^2}{\sum y^2}$$

- R_9^2: Robust version using medians to resist outliers.

$$R_9^2 = 1 - \left(\frac{M\{|y_i - \hat{y}_i|\}}{M\{|y_i - \bar{y}|\}}\right)^2$$

where M represents the median of the sample.

For degree of freedom adjustment `adjusted = TRUE`, refer to r2_adjusted.

Value

An object of class `r2_kvr2`, which is a list containing the calculated values for each R^2 formula.

Note

The power regression model must be based on a logarithmic transformation.

The auto-selection between linear regression and power regression models is determined by whether the dependent variable's name contains "log". If the name "log" is intentionally used for a linear regression model, the selection cannot be made correctly.

References

Tarald O. Kvalseth (1985) Cautionary Note about R 2 , The American Statistician, 39:4, 279-285, doi:10.1080/00031305.1985.10479448

Box, George E. P., Hunter, William G., Hunter, J. Stuart. (1978) Statistics for experimenters: an introduction to design, data analysis, and model building. New York, United States, J. Wiley, p. 462-473, ISBN:9780471093152.

See Also

print.r2_kvr2() r2_adjusted()

Examples

```
# Example data set 1. Kvalseth (1985).
df1 <- data.frame(x = c(1:6),
                  y = c(15,37,52,59,83,92))
# Linear regression model with intercept
model_intercept1 <- lm(y ~ x, df1)
# Linear regression model without intercept
model_without1 <- lm(y ~ x - 1, df1)
# Power regression model
model_power1 <- lm(log(y) ~ log(x), df1)
r2(model_intercept1)
r2(model_without1)
r2(model_power1)
# Example data set 2. Kvalseth (1985).
df2 <- data.frame(x = 6:13,
                  y = c(3882, 1266, 733, 450, 410, 305, 185, 112))
power_model2 <- lm(log((y/7343)) ~ log(x), data = df2)
r2(power_model2)
# Example of a Multiple Regression Analysis Model.
# The data for two independent variables given by Box et al. (1978, p. 462)
# as used in Kvalseth (1985).
df3 <- data.frame(x1 = c(0.34, 0.34, 0.58, 1.26, 1.26, 1.82),
                  x2 = c(0.73, 0.73, 0.69, 0.97, 0.97, 0.46),
                  y = c(5.75, 4.79, 5.44, 9.09, 8.59, 5.09))
# Multiple regression analysis model with intercept
model_intercept3 <- lm(y ~ x1 + x2, df3)
# Multiple regression analysis model without intercept
model_without3 <- lm(y ~ x1 + x2 - 1, df3)
# Multiple power regression analysis model
model_power3 <- lm(log(y) ~ log(x1) + log(x2), df3)
r2(model_intercept3)
r2(model_without3)
r2(model_power3)
```

---

r2_adjusted — *Calculate the adjusted determination coefficient*

Description

Calculate the adjusted coefficient of determination by entering the regression model and coefficient of determination. See details.

Usage

```
r2_adjusted(model, r2)
```

Arguments

`model`	A linear model or power regression model of the `lm`.
`r2`	A numeric. Coefficient of determination.

Details

The adjustment factor a is calculated using the following formula.

$$a = (n-1)/(n-k-1)$$

n is the sample size, and k is the number of parameters in the regression model.

R_a^2 ($R^2 adjusted$) is calculated using the following formula.

$$R_a^2 = 1 - a(1 - R^2)$$

This function performs freedom-of-degrees adjustment for all coefficients based on the above formula. However, Kvalseth (1985) recommends applying freedom-of-degrees adjustment only to R_1^2 and R_9^2, based on the principle of consistency in coefficients. Furthermore, there is no basis for applying the same type of adjustment to R_6^2 (the square of the correlation coefficient) or to R_7^2 and R_8^2, which depend on specific model forms.

For details on each coefficient of determination, refer to `r2()`.

Value

A numeric vector or a list of class `r2_kvr2` containing the adjusted R^2 values. Each element represents the adjusted version of the corresponding R^2 definition, accounting for the degrees of freedom.

References

Tarald O. Kvalseth (1985) Cautionary Note about R 2 , The American Statistician, 39:4, 279-285, doi:10.1080/00031305.1985.10479448

See Also

`r2()`

Index

comp_fit, 2
comp_fit(), *4*

MAE (comp_fit), 2
MSE (comp_fit), 2

print.comp_kvr2 (print.r2_kvr2), 4
print.comp_kvr2(), *3*
print.r2_kvr2, 4
print.r2_kvr2(), *7*

r2, 4
r2(), *4*, *8*
r2_1 (r2), 4
r2_2 (r2), 4
r2_3 (r2), 4
r2_4 (r2), 4
r2_5 (r2), 4
r2_6 (r2), 4
r2_7 (r2), 4
r2_8 (r2), 4
r2_9 (r2), 4
r2_adjusted, *6*, 7
r2_adjusted(), *4*, *7*
RMSE (comp_fit), 2