Package ‘LogRegEquiv’

August 13, 2021

Version 0.1.2
Date 2021-08-11
Title Logistic Regression Equivalence
Description Tools for assessing equivalence of similar Logistic Regression models.
Author Guy Ashiri-Prossner
Maintainer Guy Ashiri-Prossner <guy.ashiri@mail.huji.ac.il>
Encoding UTF-8
RoxygenNote 7.1.1
License MIT + file LICENSE
Imports stats
Suggests knitr, rmarkdown, testthat
VignetteBuilder knitr
Depends R (>= 2.10)
LazyData true
NeedsCompilation no
Repository CRAN
Date/Publication 2021-08-13 09:10:07 UTC

R topics documented:

beta_equivalence .................................................. 2
brier_score .......................................................... 2
descriptive_equiv .................................................... 3
individual_predictive_equiv ....................................... 4
performance_equiv ................................................... 4
ptg_stud_data ....................................................... 5
ptg_stud_f_test ..................................................... 6
ptg_stud_f_train ..................................................... 7
ptg_stud_m_test ..................................................... 7
ptg_stud_m_train ................................................... 8
sigmoid ............................................................... 8
beta_equivalence

**Description**

This function takes two logistic regression models $M_A, M_B$, sensitivity level $\delta_{\beta}$ and significance level $\alpha$. It checks whether the coefficient vectors are equivalent.

**Usage**

```r
beta_equivalence(model_a, model_b, delta, alpha)
```

**Arguments**

- `model_a`: logistic regression model $M_A$
- `model_b`: logistic regression model $M_B$
- `delta`: equivalence sensitivity level $\delta_{\beta}$. This could either be a scalar or a vector with length matching the number of coefficients.
- `alpha`: significance level $\alpha$

**Value**

- `equivalence`: are the coefficient vectors equivalent? (boolean)
- `test_statistic`: Equivalence test statistic
- `critical value`: a level-$\alpha$ critical value
- `ncp`: non-centrality parameter
- `p_value`: P-value

brier_score

**Description**

This function takes a observations vector $y$ and matching predictions vector $\pi$. It returns the Brier score for the predictions. Unless specified otherwise, input containing NAs will result with an NA.

**Usage**

```r
brier_score(y, pi, na.rm = FALSE)
```
Arguments

- \( y \) the observations vector
- \( \pi \) the predictions vector
- \texttt{na.rm} ignore NA? (optional)

Value

The Brier score \( \frac{1}{N} \sum_{i=1}^{N} (y_i - \pi_i)^2 \)

Examples

\[
\text{brier_score(rbinom(10, 1, seq(0.1, 1, 0.1)), seq(0.1, 1, 0.1))}
\]

---

Description

This function takes two datasets \( X_A, X_B \), regression formula, significance level \( \alpha \) and sensitivity level \( \delta_B \) (either vector or scalar). It builds a logistic regression model for each of the datasets and then checks whether the obtained coefficient vectors are equivalent, using the \texttt{beta_equivalence} function.

Usage

\[
\text{descriptive_equiv(data_a, data_b, formula, delta, alpha = 0.05)}
\]

Arguments

- \texttt{data_a} dataset \( X_A \) for model \( M_A \)
- \texttt{data_b} dataset \( X_B \) for model \( M_B \)
- \texttt{formula} logistic regression formula
- \texttt{delta} equivalence sensitivity level \( \delta_B \)
- \texttt{alpha} significance level \( \alpha \) (defaults to 0.05)

Value

- \texttt{equivalence} the \texttt{beta_equivalence} function output
- \texttt{model_a} logistic regression model \( M_A \)
- \texttt{model_b} logistic regression model \( M_B \)
individual_predictive_equiv

Description

This function takes two logistic regression models $M_A, M_B$, test data, significance level $\alpha$ and allowed flips ratio $r$. It checks whether the models produce equivalent log-odds for the given test set and returns various figures.

Usage

individual_predictive_equiv(model_a, model_b, test_data, r = 0.1, alpha = 0.05)

Arguments

- model_a: logistic regression model $M_A$
- model_b: logistic regression model $M_B$
- test_data: testing dataset
- r: ratio of allowed 'flips' (defaults to 0.1)
- alpha: significance level $\alpha$ (defaults to 0.05)

Value

- equivalence: Are models $M_A, M_B$ producing equivalent log-odds for the given test data? (boolean)
- test_statistic: The test statistic
- critical_value: $\alpha$ level critical value the test
- xi_bar: Mean $\xi$ value for the test
- delta_theta: Calculated equivalence parameter
- p_value: P-value

performance_equiv

Description

This function takes two logistic regression models $M_A, M_B$, test data, significance level $\alpha$ and sensitivity level $\delta_B$. It checks whether the models perform equivalently on the test set and returns various figures.

Usage

performance_equiv(model_a, model_b, test_data, dv_index, t = 0.1, alpha = 0.05)
Arguments

- **model_a**: logistic regression model $M_A$
- **model_b**: logistic regression model $M_B$
- **test_data**: testing dataset
- **dv_index**: column number of the dependent variable
- **t**: acceptable tolerance level (defaults to 0.1)
- **alpha**: significance level $\alpha$ (defaults to 0.05)

Value

- **equivalence**: Are models $M_A, M_B$ producing equivalent Brier scores for the given test data? (boolean)
- **brier_score_ac** $M_A$ Brier score on the testing data
- **brier_score_bc** $M_B$ Brier score on the testing data
- **diff_sd**: SD of the Brier differences
- **test_stat_l** $t_L$ equivalence boundary for the test
- **test_stat_u** $t_U$ equivalence boundary for the test
- **crit_val**: a level-$\alpha$ critical value for the test
- **delta_B**: Calculated equivalence parameter
- **p_value_l**: P-value for $t_L$
- **p_value_u**: P-value for $t_U$

---

**ptg_stud_data**

*Student Performance Data Set*

Description

Data from a student achievement in secondary education of two Portuguese schools. Full attribute description could be found in the source webpage.

Usage

- **ptg_stud_data**

Format

An object of class `data.frame` with 649 rows and 31 columns.
Details

The data used is taken from the Student Performance Data. The original data consists of 30 covariates (13 binary, 11 ordinal, 4 categorical, 2 numerical) and a numerical output variable indicating the students final grade in Portuguese Language course.

The data was split by gender (F/M) \( n_f = 383, n_m = 266 \). The target variable G3 was converted to binary, final_fail which indicates the cases where G3 < 10.

Next, each sub-population was divided into training and testing data, using a 4:1 ratio.

Source


References


See Also


An object of class data.frame with 77 rows and 30 columns.

See Also

ptg_stud_data
**ptg_stud_f_train**

*Student Performance Data Set - female training data*

**Description**

Student Performance Data Set - female training data

**Usage**

`ptg_stud_f_train`

**Format**

An object of class `data.frame` with 306 rows and 30 columns.

**See Also**

`ptg_stud_data`

---

**ptg_stud_m_test**

*Student Performance Data Set - male testing data*

**Description**

Student Performance Data Set - male testing data

**Usage**

`ptg_stud_m_test`

**Format**

An object of class `data.frame` with 53 rows and 30 columns.

**See Also**

`ptg_stud_data`
sigmoid

Description

This function takes a number \( \theta \) and returns its respective sigmoid probability \( \frac{e^{\theta}}{1+e^{\theta}} \). This is used in logistic regression to model \( P(y = 1|x) \).

Usage

sigmoid(theta)

Arguments

theta the linear predictor

Value

the sigmoid probability

Examples

sigmoid(0)

ptg_stud_m_train Student Performance Data Set - male training data

Description

Student Performance Data Set - male training data

Usage

ptg_stud_m_train

Format

An object of class data.frame with 213 rows and 30 columns.

See Also

ptg_stud_data
Index

* Brier-score
  brier_score, 2

* datasets
  ptg_stud_data, 5
  ptg_stud_f_test, 6
  ptg_stud_f_train, 7
  ptg_stud_m_test, 7
  ptg_stud_m_train, 8

beta_equivalence, 2
brier_score, 2

descriptive_equiv, 3

individual_predictive_equiv, 4

performance_equiv, 4
ptg_stud_data, 5
ptg_stud_f_test, 6
ptg_stud_f_train, 7
ptg_stud_m_test, 7
ptg_stud_m_train, 8

sigmoid, 8