

Package ‘flexOR’

April 16, 2024

Encoding UTF-8

Type Package

Version 1.0.0

Date 2024-03-25

Title Flexible Odds Ratio Curves

Maintainer Marta Azevedo <marta.vasconcelos4@gmail.com>

Depends R (>= 3.1.0)

Imports gam, graphics, grDevices, mgcv, stats

Suggests knitr, mlbench, R.rsp, rmarkdown, plotly

Description Provides flexible odds ratio curves that enable modeling non-linear relationships between continuous predictors and binary outcomes. This package facilitates a deeper understanding of the impact of each continuous predictor on the outcome by presenting results in terms of odds ratio (OR) curves based on splines. These curves allow for comparison against a specified reference value, aiding in the interpretation of the predictor's effect.

License GPL-3

URL <https://github.com/martaaaa/flexOR>

BugReports <https://github.com/martaaaa/flexOR/issues>

LazyLoad yes

NeedsCompilation no

RoxygenNote 7.3.1

VignetteBuilder knitr

Author Marta Azevedo [aut, cre] (<<https://orcid.org/0009-0003-4412-6991>>),
Luis Meira-Machado [aut] (<<https://orcid.org/0000-0002-8577-7665>>),
Artur Araujo [aut] (<<https://orcid.org/0000-0003-1419-4208>>)

Repository CRAN

Date/Publication 2024-04-16 13:40:02 UTC

R topics documented:

flexOR-package	2
AICc	3
dfgam	4
flexOR	5
floor_to	6
plot.OR	7
predict.OR	9

Index	11
--------------	-----------

flexOR-package	<i>flexOR: Flexible Odds Ratio Curves</i>
----------------	---

Description

Provides flexible odds ratio curves that enable modeling non-linear relationships between continuous predictors and binary outcomes. This package facilitates a deeper understanding of the impact of each continuous predictor on the outcome by presenting results in terms of odds ratio (OR) curves based on splines. These curves allow for comparison against a specified reference value, aiding in the interpretation of the predictor's effect.

Author(s)

Maintainer: Marta Azevedo <marta.vasconcelos4@gmail.com> ([ORCID](#))

Authors:

- Luis Meira-Machado <lmachado@math.uminho.pt> ([ORCID](#))
- Artur Araujo <artur.stat@gmail.com> ([ORCID](#))

References

- Hosmer, D. W. and Lemeshow, S. and Sturdivant, R. X. (2013). *Applied Logistic Regression: Third Edition*, John Wiley and Sons Inc., New York, NY.
- Royston, P. and Altman, D. G. and Sauerbrei, W. (2006). Dichotomizing continuous predictors in multiple regression: A bad idea. *Statistics in Medicine*, **25**(1), 127–141. doi:[10.1002/sim.2331](#)
- Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, Chapman & Hall/CRC, New York, NY.
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R: Second Edition*, Chapman & Hall/CRC, London, UK.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6), 716–723. doi:[10.1109/TAC.1974.1100705](#)
- Hurvich, C. M. and Simonoff, J. S. and Tsai, C. (1998). Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **60**(2), 271–293. doi:[10.1111/14679868.00125](#)

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**(2), 461–464. doi:10.1214/aos/1176344136

Cadarso-Suárez, C. and Meira-Machado, L. and Kneib, T. and Gude, F. (2010). Flexible hazard ratio curves for continuous predictors in multi-state models: an application to breast cancer data. *Statistical Modelling*, **10**(3), 291–314. doi:10.1177/1471082X0801000303

Meira-Machado, L. and Cadarso-Suárez, C. and Gude, F. and Araújo, A. (2013). smoothHR: An R Package for Pointwise Nonparametric Estimation of Hazard Ratio Curves of Continuous Predictors. *Computational and Mathematical Methods in Medicine*, **2013**, 11 pages. doi:10.1155/2013/745742

de Boor, C. (2001). *A Practical Guide to Splines: Revised Edition*, Springer, New York, NY.

Wood, S. N. and Pya, N. and Säfken, B. (2016). Smoothing Parameter and Model Selection for General Smooth Models. *Journal of the American Statistical Association*, **111**(516), 1548-1563. doi:10.1080/01621459.2016.1180986

See Also

Useful links:

- <https://github.com/martaanaa/flexOR>
- Report bugs at <https://github.com/martaanaa/flexOR/issues>

AICc

AICc: Calculation for GAM Models

Description

Calculates AICc (Akaike Information Criterion corrected for small sample sizes) for Generalized Additive Models (GAM).

Usage

AICc(object)

Arguments

object An object of class "Gam" or "gam" representing a fitted GAM model.

Details

This function calculates the AICc value (Akaike Information Criterion corrected for small sample sizes) for a given GAM model. AICc is a measure of model fit that penalizes the number of parameters in the model to avoid overfitting.

Value

A numeric value representing the AICc for the GAM model.

See Also

[gam](#), [logLik](#), [AIC](#)

Examples

```
# Load dataset
data(PimaIndiansDiabetes2, package="mlbench");

# Fit GAM model
fit <- mgcv::gam(
  diabetes ~ s(age) + s(mass) + s(pedigree) + pressure + glucose,
  data=PimaIndiansDiabetes2,
  family=binomial
);

# Calculate AICc
AICc(fit);
```

dfgam

dfgam: Degrees of Freedom Selection for GAM Models

Description

Computes the degrees of freedom for specified non-linear predictors in a GAM model. The user can choose between AIC (Akaike Information Criterion), AICc (AIC corrected for small sample sizes), or BIC (Bayesian Information Criterion) as the selection criteria. This function is useful for determining the appropriate degrees of freedom for smoothing terms in GAMs.

Usage

```
dfgam(
  response,
  nl.predictors,
  other.predictors = NULL,
  smoother = "s",
  method = "AIC",
  data,
  step = NULL
)
```

Arguments

`response` The response variable as a formula.

`nl.predictors` A character vector specifying the non-linear predictors.

`other.predictors` A character vector specifying other predictors if needed.

smoother	The type of smoothing term, currently only "s" is supported.
method	The selection method, one of "AIC", "AICc", or "BIC".
data	The data frame containing the variables.
step	The step size for grid search when there are multiple non-linear predictors.

Value

A list containing the following components:

- `fit`: The fitted GAM model.
- `df`: A numeric vector of degrees of freedom for each non-linear predictor.
- `method`: The selection method used (AIC, AICc, or BIC).
- `nl.predictors`: The non-linear predictors used in the model.
- `other.predictors`: Other predictors used in the model if specified.

Examples

```
# Load dataset
library(gam)
data(PimaIndiansDiabetes2, package="mlbench");

# Calculate degrees of freedom using AIC
df2 <- dfgam(
  response="diabetes",
  nl.predictors=c("age", "mass"),
  other.predictors=c("pedigree"),
  smoother="s",
  method="AIC",
  data=PimaIndiansDiabetes2
);

print(df2$df);
```

flexOR

flexOR: Flexible Odds Ratio Computation for GAM Models

Description

Computes odds ratios for predictors in GAM models. It provides flexibility in specifying predictors using either a data frame, a response variable, and a formula or a pre-fitted GAM model. The function is useful for understanding the impact of predictors on binary outcomes in GAMs.

Usage

```
flexOR(data, response = NULL, formula = NULL)
```

Arguments

data	A data frame containing the variables.
response	The response variable as a character string.
formula	A formula specifying the model.

Details

It accepts two different ways of specifying the model: by providing the data frame and response variable or by specifying the formula.

Value

A list containing the following components:

- `dataset`: The dataset used for the analysis.
- `formula`: The formula used in the GAM model.
- `gamfit`: The fitted GAM model.
- `response`: The response variable used in the analysis.

Examples

```
library(gam);

# Load dataset
data(PimaIndiansDiabetes2, package="mlbench");

# Calculate odds ratios using flexOR
df_result <- flexOR(data = PimaIndiansDiabetes2, response = "diabetes",
formula=~ s(age) + s(mass) + s(pedigree) + pressure + glucose)
print(df_result)
```

floor_to

floor_to: Rounds a number down

Description

Takes a numeric value or vector (x) and rounds it down to the nearest multiple of a specified base (to). If the `to` argument is not provided, it defaults to rounding down to the nearest integer. The result is returned as a numeric value or vector of the same length as the input.

Usage

```
floor_to(x, to = 1)
```

Arguments

x numeric value or vector that you want to round down or floor.
to the base to which you want to round down the input value(s). The default value is set to 1, meaning the function will round down to the nearest integer.

Value

The number rounded down to the specified multiple.

Examples

```
floor_to(7, 3); # Returns 6, as 6 is the largest multiple of 3 less than or equal to 7.  
floor_to(5, 2); # Returns 4, as 4 is the largest multiple of 2 less than or equal to 5.
```

plot.OR

plot.OR: Plot Smooth Odds Ratios

Description

Plots smooth odds ratios along with confidence intervals for a specified predictor.

Usage

```
## S3 method for class 'OR'  
plot(  
  x,  
  predictor,  
  prob = NULL,  
  ref.value = NULL,  
  conf.level = 0.95,  
  round.x = NULL,  
  ref.label = NULL,  
  col,  
  col.area,  
  main,  
  xlab,  
  ylab,  
  lty,  
  xlim,  
  ylim,  
  xx,  
  ylog = TRUE,  
  log = ifelse(ylog, "", "y"),  
  ...  
)
```

Arguments

x	An object of class "OR" generated by the <code>flexOR</code> function.
predictor	The name of the predictor variable for which to plot the smooth odds ratios.
prob	The probability level for the confidence interval. Default is NULL.
ref.value	The predicted value at which to calculate the smooth odds ratios. Default is NULL.
conf.level	The confidence level for the intervals. Default is 0.95.
round.x	The number of decimal places to round the predictor variable values. Default is NULL.
ref.label	The label for the reference value of the predictor variable. Default is NULL.
col	Vector of colors for plotting. Default is <code>c("black", "black", "grey85")</code> .
col.area	Vector of colors for the confidence intervals.
main	The title of the plot. Default is generated based on the predictor variable.
xlab	Label for the x-axis. Default is the name of the predictor variable.
ylab	Label for the y-axis. Default is "Ln OR(Z,Zref)" if logarithmic scale is used, else "OR(Z,Zref)".
lty	Vector of line types for plotting. Default is <code>c(1, 3)</code> .
xlim	Range of the x-axis. Default is NULL.
ylim	Range of the y-axis. Default is NULL.
xx	Values for tick marks on the x-axis. Default is NULL.
ylog	Logical. If TRUE, y-axis is on a logarithmic scale. Default is TRUE.
log	Use a logarithmic scale for the y-axis (alternative argument name).
...	Additional arguments passed to plotting functions.

Value

This function doesn't return a value. It is used for generating a plot.

Examples

```
library(gam);

# Load dataset
data(PimaIndiansDiabetes2, package="mlbench");

mod1 <- flexOR(
  data=PimaIndiansDiabetes2,
  response="diabetes",
  formula=~s(age, 3.3) + s(mass, 4.1) + pedigree
);
```

predict.OR

predict.OR : Predict Method for OR Objects

Description

Predicts values using a fitted OR model.

Usage

```
## S3 method for class 'OR'
predict(
  object,
  predictor,
  prob = NULL,
  ref.value = NULL,
  conf.level = 0.95,
  prediction.values = NULL,
  round.x = NULL,
  ref.label = NULL,
  ...
)
```

Arguments

<code>object</code>	An object of class "OR."
<code>predictor</code>	The predictor variable for which you want to make predictions.
<code>prob</code>	Probability value for prediction. Use 0 for point prediction, 0.5 for median, or a custom value between 0 and 1.
<code>ref.value</code>	Optional custom prediction value (use with <code>prob=NULL</code>).
<code>conf.level</code>	Confidence level for prediction intervals (default is 0.95).
<code>prediction.values</code>	Vector of specific prediction values to calculate.
<code>round.x</code>	Number of decimal places to round the prediction values (default is 5).
<code>ref.label</code>	Label for the predictor variable in the output (optional).
<code>...</code>	Additional arguments (not used in this function).

Details

This function predicts values and prediction intervals using a fitted OR model.

Value

A matrix with predicted values and prediction intervals.

Examples

```
library(gam);

# Load the Pima Indians Diabetes dataset
data(PimaIndiansDiabetes2, package="mlbench");

# Calculate smooth odds ratios using flexOR
mod1 <- flexOR(
  data=PimaIndiansDiabetes2,
  response="diabetes",
  formula= ~ s(age) + s(mass) + s(pedigree) + pressure + glucose
);

# Predict the probabilities using predict.OR
predict(mod1, predictor="age", ref.value=40)
```

Index

- * **arith**
 - floor_to, 6
 - * **hplot**
 - plot.OR, 7
 - * **methods**
 - plot.OR, 7
 - predict.OR, 9
 - * **models**
 - AICc, 3
 - dfgam, 4
 - flexOR, 5
 - plot.OR, 7
 - predict.OR, 9
 - * **nonlinear**
 - AICc, 3
 - dfgam, 4
 - flexOR, 5
 - plot.OR, 7
 - predict.OR, 9
 - * **package**
 - flexOR-package, 2
 - * **regression**
 - AICc, 3
 - dfgam, 4
 - flexOR, 5
 - plot.OR, 7
 - predict.OR, 9
 - * **smooth**
 - AICc, 3
 - dfgam, 4
 - flexOR, 5
 - plot.OR, 7
 - predict.OR, 9
- AIC, 4
- AICc, 3
- dfgam, 4
- flexOR, 5, 8
- flexOR-package, 2
- floor_to, 6
- gam, 4
- logLik, 4
- plot.OR, 7
- predict.OR, 9