

# Package ‘EffectStars’

October 9, 2017

**Type** Package

**Title** Visualization of Categorical Response Models

**Version** 1.8

**Date** 2017-10-09

**Depends** VGAM

**Author** Gunther Schauburger

**Maintainer** Gunther Schauburger <gunther.schauburger@stat.uni-muenchen.de>

**Description** Notice: The package EffectStars2 provides a more up-to-date implementation of effect stars! EffectStars provides functions to visualize regression models with categorical response. The effects of the variables are plotted with star plots in order to allow for an optical impression of the fitted model.

**License** GPL-2

**LazyLoad** yes

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2017-10-09 09:15:01 UTC

## R topics documented:

alligator	2
BEPS	3
coffee	4
EffectStars	5
election	5
insolvency	8
PID	9
plebiscite	10
star.cumulative	11
star.nominal	14
star.sequential	19
womenlabour	22

<b>Index</b>	<b>24</b>
--------------	-----------

---

`alligator`*Alligator Food*

---

**Description**

The data describe the food choice of alligators, they originate from a study of the Florida Game and Fresh Water Commission.

**Usage**

```
data(alligator)
```

**Format**

A data frame with 219 observations on the following 4 variables.

Food Food type with levels bird, fish, invert, other and rep

Size Size of the alligator with levels <2.3 and >2.3

Gender Gender with levels female and male

Lake Name of the lake with levels George, Hancock, Oklawaha and Trafford

**Source**

<http://www.stat.ufl.edu/~aa/cda/sas/sas.html>

**References**

Agresti (2002): *Categorical Data Analysis*, Wiley.

**Examples**

```
## Not run:  
data(alligator)  
star.nominal(Food ~ Size + Lake + Gender, data = alligator, nlines = 2)  
  
## End(Not run)
```

---

BEPS

*British Election Panel Study*

---

### Description

These data are drawn from the 1997-2001 British Election Panel Study (BEPS).

### Usage

```
data(BEPS)
```

### Format

A data frame with 1525 observations on the following 10 variables.

Europe An 11-point scale that measures respondents' attitudes toward European integration. High scores represent eurosceptic sentiment

Leader\_Cons Assessment of the Conservative leader Hague, 1 to 5

Leader\_Labour Assessment of the Labour leader Blair, 1 to 5

Leader\_Liberals Assessment of the Liberals leader Kennedy, 1 to 5

Vote Party Choice with levels Conservative, Labour and Liberal Democrat

Age Age in years

Gender Gender with levels female and male

Political\_Knowledge Knowledge of parties' positions on European integration, 0 to 3

National\_Economy Assessment of current national economic conditions, 1 to 5

Household Assessment of current household economic conditions, 1 to 5

### Source

R package carData: [BEPS](#)

### References

British Election Panel Study (BEPS)

J. Fox and R. Andersen (2006): *Effect displays for multinomial and proportional-odds logit models*. *Sociological Methodology* 36, 225–255

### Examples

```
## Not run:  
data(BEPS)
```

```
BEPS$Europe<-scale(BEPS$Europe)  
BEPS$Age<-scale(BEPS$Age)  
BEPS$Leader_Labour<-BEPS$Leader_Labour-BEPS$Leader_Cons  
BEPS$Leader<-BEPS$Leader_Labour
```

```

BEPS$Leader_Liberals<-BEPS$Leader_Liberals-BEPS$Leader_Cons

star.nominal(Vote ~ Age + Household + National_Economy + Household + Leader +
Europe + Political_Knowledge + Gender, data = BEPS,
xij = list(Leader~Leader_Labour+Leader_Liberals), catstar = FALSE, symmetric = FALSE)

## End(Not run)

```

---

coffee

*Coffee Brands*

---

### Description

The data frame is part of a long-term panel about the choice of coffee brands in 2111 households. The explanatory variables either refer to the household as a whole or to the head of the household.

### Usage

```
data(coffee)
```

### Format

A data frame with 2111 observations on the following 8 variables.

Education Educational level with levels no Highschool and Highschool

PriceSensitivity Price sensitivity with levels not sensitive and sensitive

Income Income with levels < 2499 and >= 2500

SocialLevel Social level with levels high and low

Age Age with levels < 49 and >= 50

Brand Coffee Brand with levels Jacobs, JacobsSpecial, Aldi, AldiSpecial, Eduscho, EduschoSpecial, Tchibo, TchiboSpecial and Others

Amount Amount of packs with levels 1 and >= 2

Persons Number of persons in household

### References

Gesellschaft für Konsumforschung (GfK)

### Examples

```

## Not run:
data(coffee)

star.nominal(Brand ~ Amount + Age + SocialLevel + Income + Persons +
              PriceSensitivity + Education, coffee, cex.cat = 0.5, cex.labels = 0.8)

## End(Not run)

```

**Description**

**The package EffectStars2 provides a more up-to-date implementation of effect stars!**

The package provides functions that visualize categorical regression models.

Included models are the multinomial logit model, the sequential logit model and the cumulative logit model.

The exponentials of the effects of the predictors are plotted as star plots showing the strengths of the effects.

In addition p-values for the effect of predictors are given.

Various data sets and examples are provided.

The plots should in general be exported to file formats like pdf, ps or png to receive the optimal display. Plotting in R devices may not provide the optimal results.

For further details see [star.nominal](#), [star.sequential](#) and [star.cumulative](#).

**Author(s)**

Gunther Schauburger

<gunther@stat.uni-muenchen.de>

<http://www.semsto.statistik.uni-muenchen.de/personen/doktoranden/schauburger/index.html>

**References**

Tutz, G. and Schauburger, G. (2012): *Visualization of Categorical Response Models - from Data Glyphs to Parameter Glyphs*, Journal of Computational and Graphical Statistics 22(1), 156-177.

Gerhard Tutz (2012): *Regression for Categorical Data*, Cambridge University Press

**See Also**

[star.nominal](#), [star.sequential](#), [star.cumulative](#)

**Description**

The data set contains data from the German Longitudinal Election Study. The Response Categories refer to the five dominant parties in Germany. The explanatory variables refer to the declarations of single voters.

**Usage**

```
data(election)
```

**Format**

A data frame with 816 observations on the following 30 variables.

Age Standardized age of the voter

AgeOrig Unstandardized age of the voter

Partychoice Party Choice with levels CDU, SPD, FDP, Greens and Left Party

Gender Gender with levels female and male

West Regional provenance (West-Germany or East-Germany) with levels east and west

Union Member of a Union with levels no member and member

Highschool Educational level with levels no highschool and highschool

Unemployment Unemployment with levels not unemployed and unemployed

Pol.Interest Political Interest with levels very interested and less interested

Democracy Satisfaction with the functioning of democracy with levels satisfied and not satisfied

Religion Religion with levels evangelical, catholic and other religion

Social\_CDU Difference in attitude towards the socioeconomic dimension of politics between respondent and CDU

Social\_SPD Difference in attitude towards the socioeconomic dimension of politics between respondent and SPD

Social\_FDP Difference in attitude towards the socioeconomic dimension of politics between respondent and FDP

Social\_Greens Difference in attitude towards the socioeconomic dimension of politics between respondent and the Greens

Social\_Left Difference in attitude towards the socioeconomic dimension of politics between respondent and the Left party

Immigration\_CDU Difference in attitude towards immigration of foreigners between respondent and CDU

Immigration\_SPD Difference in attitude towards immigration of foreigners between respondent and SPD

Immigration\_FDP Difference in attitude towards immigration of foreigners between respondent and FDP

Immigration\_Greens Difference in attitude towards immigration of foreigners between respondent and the Greens

Immigration\_Left Difference in attitude towards immigration of foreigners between respondent and the Left party

Nuclear\_CDU Difference in attitude towards nuclear energy between respondent and CDU

Nuclear\_SPD Difference in attitude towards nuclear energy between respondent and SPD

Nuclear\_FDP Difference in attitude towards nuclear energy between respondent and FDP

Nuclear\_Greens Difference in attitude towards nuclear energy between respondent and the Greens  
 Nuclear\_Left Difference in attitude towards nuclear energy between respondent and the Left party  
 Left\_Right\_CDU Difference in attitude towards the positioning on a political left-right scale between respondent and CDU  
 Left\_Right\_SPD Difference in attitude towards the positioning on a political left-right scale between respondent and SPD  
 Left\_Right\_FDP Difference in attitude towards the positioning on a political left-right scale between respondent and FDP  
 Left\_Right\_Greens Difference in attitude towards the positioning on a political left-right scale between respondent and the Greens  
 Left\_Right\_Left Difference in attitude towards the positioning on a political left-right scale between respondent and the Left party

## References

German Longitudinal Election Study (GLES)

## Examples

```
## Not run:
data(election)

# simple multinomial logit model
star.nominal(Partychoice ~ Age + Religion + Democracy + Pol.Interest +
              Unemployment + Highschool + Union + West + Gender, election)

# Use effect coding for the categorical predictor religion
star.nominal(Partychoice ~ Age + Religion + Democracy + Pol.Interest +
              Unemployment + Highschool + Union + West + Gender, election,
              pred.coding = "effect")

# Use reference category "FDP" instead of symmetric side constraints
star.nominal(Partychoice ~ Age + Religion + Democracy + Pol.Interest +
              Unemployment + Highschool + Union + West + Gender, election,
              refLevel = 3, symmetric = FALSE)

# Use category-specific covariates, subtract values for reference
# category CDU
election[,13:16] <- election[,13:16] - election[,12]
election[,18:21] <- election[,18:21] - election[,17]
election[,23:26] <- election[,23:26] - election[,22]
election[,28:31] <- election[,28:31] - election[,27]

election$Social <- election$Social_SPD
election$Immigration <- election$Immigration_SPD
election$Nuclear <- election$Nuclear_SPD
election$Left_Right <- election$Left_Right_SPD

star.nominal(Partychoice ~ Social + Immigration + Nuclear + Left_Right + Age +
              Religion + Democracy + Pol.Interest + Unemployment + Highschool + Union + West +
```

```

Gender, data = election,
xij = list(Social ~ Social_SPD + Social_FDP + Social_Greens + Social_Left,
Immigration ~ Immigration_SPD + Immigration_FDP + Immigration_Greens + Immigration_Left,
Nuclear ~ Nuclear_SPD + Nuclear_FDP + Nuclear_Greens + Nuclear_Left,
Left_Right ~ Left_Right_SPD + Left_Right_FDP + Left_Right_Greens + Left_Right_Left),
symmetric = FALSE)

## End(Not run)

```

---

insolvency

*Insolvency data*


---

### Description

The data set originates from the Munich founder study. The data were collected on business founders who registered their new companies at the local chambers of commerce in Munich and surrounding administrative districts. The focus was on survival of firms measured in 7 categories, the first six represent failure in intervals of six months, the last category represents survival time beyond 36 months.

### Usage

```
data(insolvency)
```

### Format

A data frame with 1224 observations on the following 16 variables.

Insolvency Survival of firms in ordered categories with levels 1 < 2 < 3 < 4 < 5 < 6 < 7

Sector Economic Sector with levels industry, commerce and service industry

Legal Legal form with levels small trade, one man business, GmbH and GbR, KG, OHG

Location Location with levels residential area and business area

New\_Foundation New Foundation or take-over with levels new foundation and take-over

Pecuniary\_Reward Pecuniary reward with levels main and additional

Seed\_Capital Seed capital with levels < 25000 and > 25000

Equity\_Capital Equity capital with levels no and yes

Debt\_Capital Debt capital with levels no and yes

Market Market with levels local and national

Clientele Clientele with levels wide spread and small

Degree Educational level with levels no A-levels and A-Levels

Gender Gender with levels female and male

Experience Professional experience with levels < 10 years and > 10 years

Employees Number of employees with levels 0 or 1 and > 2

Age Age of the founder at formation of the company



**Source**

Münchener Gründer Studie

**References**

Brüderl, J. and Preisendörfer, P. and Ziegler, R. (1996): *Der Erfolg neugegründeter Betriebe: eine empirische Studie zu den Chancen und Risiken von Unternehmensgründungen*, Duncker & Humblot.

**Examples**

```
## Not run:
data(insolvency)

star.sequential(Insolvency ~ Sector + Legal + Pecuniary_Reward + Seed_Capital
+ Debt_Capital + Employees, insolvency, test.glob = FALSE, globcircle = TRUE, dist.x = 1.3)

star.cumulative(Insolvency ~ Sector + Employees, insolvency, select = 2:4)

## End(Not run)
```

---

 PID

*Party Identification*


---

**Description**

Subset of the 1996 American National Election Study.

**Usage**

```
data(election)
```

**Format**

A data frame with 944 observations on the following 6 variables.

TVnews Days in the past week spent watching news on TV

PID Party identification with levels Democrat, Independent and Republican

Income Income

Education Educational level with levels low (no college) and high (at least college)

Age Age in years

Population Population of respondent's location in 1000s of people

**Source**

R package faraway: [nes96](#)

## Examples

```
## Not run:
data(PID)
PID$TVnews <- scale(PID$TVnews)
PID$Income <- scale(PID$Income)
PID$Age <- scale(PID$Age)
PID$Population <- scale(PID$Population)

star.nominal(PID ~ TVnews + Income + Population + Age + Education, data = PID)

## End(Not run)
```

---

plebiscite

*Chilean Plebiscite*

---

## Description

The data origin from a survey referring to the plebiscite in Chile 1988. The Chilean people had to decide, whether Augusto Pinochet would remain president for another ten years (voting yes) or if there would be presidential elections in 1989 (voting no).

## Usage

```
data(plebiscite)
```

## Format

A data frame with 2431 observations on the following 7 variables.

Gender Gender with levels female and male

Education Educational level with levels low and high

SantiagoCity Respondent from Santiago City with levels no and yes

Income Monthly Income in Pesos

Population Population size of respondent's community

Age Age in years

Vote Response with levels Abstention, No, Undecided and Yes

## Source

R package car: [Chile](#)

## References

Personal communication from FLACSO/Chile.

Fox, J. (2008): *Applied Regression Analysis and Generalized Linear Models*, Second Edition.

## Examples

```
## Not run:
data(plebiscite)
plebiscite$Population <- scale(plebiscite$Population)
plebiscite$Age <- scale(plebiscite$Age)
plebiscite$Income <- scale(plebiscite$Income)

star.nominal(Vote ~ SantiagoCity + Population + Gender + Age + Education +
Income, data = plebiscite)

## End(Not run)
```

---

star.cumulative      *Effect stars for cumulative logit models*

---

## Description

**The package EffectStars2 provides a more up-to-date implementation of effect stars!**

The function computes and visualizes cumulative logit models. The computation is done with help of the package **VGAM**. The visualization is based on the function [stars](#) from the package **graphics**.

## Usage

```
star.cumulative(formula, data, global = NULL, test.rel = TRUE, test.glob = FALSE,
partial = FALSE, globcircle = FALSE, maxit = 100, scale = TRUE,
nlines = NULL, select = NULL, dist.x = 1, dist.y = 1, dist.cov = 1,
dist.cat = 1, xpd = TRUE, main = "", col.fill = "gray90",
col.circle = "black", lwd.circle = 1, lty.circle = "longdash",
col.global = "black", lwd.global = 1, lty.global = "dotdash", cex.labels = 1,
cex.cat = 0.8, xlim = NULL, ylim = NULL)
```

## Arguments

formula	An object of class “formula”. Formula for the cumulative logit model to be fitted and visualized.
data	An object of class “data.frame” containing the covariates used in formula.
global	Numeric vector to choose a subset of predictors to be included with global coefficients. Default is to include all coefficients category-specific. Numbers refer to total amount of predictors, including intercept and dummy variables.
test.rel	Provides a Likelihood-Ratio-Test to test the relevance of the explanatory covariates. The corresponding p-values will be printed as p-rel. test.rel=FALSE might save a lot of time. See also Details.
test.glob	Provides a Likelihood-Ratio-Test to test if a covariate has to be included as a category-specific covariate (in contrast to being global). The corresponding p-values will be printed as p-global. test.glob=FALSE and globcircle=FALSE might save a lot of time. See also Details.

partial	If partial=TRUE, partial proportional odds models with only one category-specific covariate are fitted. The resulting effects of the (sub)models are plotted. For further information see Details.
globcircle	If TRUE, additional circles that represent the global effects of the covariates are plotted. test.glob=FALSE and globcircle=FALSE might save a lot of time.
maxit	Maximal number of iterations to fit the cumulative logit model. See also <a href="#">vglm.control</a> .
scale	If TRUE, the stars are scaled to equal maximal ray length.
nlines	If specified, nlines gives the number of lines in which the effect stars are plotted.
select	Numeric vector to choose only a subset of the stars to be plotted. Default is to plot all stars. Numbers refer to total amount of predictors, including intercept and dummy variables.
dist.x	Optional factor to increase/decrease distances between the centers of the stars on the x-axis. Values greater than 1 increase, values smaller than 1 decrease the distances.
dist.y	Optional factor to increase/decrease distances between the centers of the stars on the y-axis. Values greater than 1 increase, values smaller than 1 decrease the distances.
dist.cov	Optional factor to increase/decrease distances between the stars and the covariates labels above the stars. Values greater than 1 increase, values smaller than 1 decrease the distances.
dist.cat	Optional factor to increase/decrease distances between the stars and the category labels around the stars. Values greater than 1 increase, values smaller than 1 decrease the distances.
xpd	If FALSE, all plotting is clipped to the plot region, if TRUE, all plotting is clipped to the figure region, and if NA, all plotting is clipped to the device region. See also <a href="#">par</a> .
main	An overall title for the plot. See also <a href="#">plot</a> .
col.fill	Color of background of the circle. See also col in <a href="#">par</a> .
col.circle	Color of margin of the circle. See also col in <a href="#">par</a> .
lwd.circle	Line width of the circle. See also lwd in <a href="#">par</a> .
lty.circle	Line type of the circle. See also lty in <a href="#">par</a> .
col.global	Color of margin of the global effects circle. See also col in <a href="#">par</a> . Ignored, if globcircle = FALSE.
lwd.global	Line width of the global effects circle. See also lwd in <a href="#">par</a> . Ignored, if globcircle = FALSE.
lty.global	Line type of the global effects circle. See also lty in <a href="#">par</a> . Ignored, if globcircle = FALSE.
cex.labels	Size of labels for covariates placed above the corresponding star. See also cex in <a href="#">par</a> .
cex.cat	Size of labels for categories placed around the corresponding star. See also cex in <a href="#">par</a> .
xlim	Optional specification of the x coordinates ranges. See also xlim in <a href="#">plot.window</a>
ylim	Optional specification of the y coordinates ranges. See also ylim in <a href="#">plot.window</a>

## Details

The underlying models are fitted with the function `vglm` from the package **VGAM**. The family argument for `vglm` is `cumulative(parallel=FALSE)`.

The stars show the exponentials of the estimated coefficients. In cumulative logit models the exponential coefficients can be interpreted as odds. More precisely, the exponential  $e^{\gamma_{rj}}$ ,  $r = 1, \dots, k-1$  represents the multiplicative effect of the covariate  $j$  on the cumulative odds  $\frac{P(Y \leq r|x)}{P(Y > r|x)}$  if  $x_j$  increases by one unit.

In addition to the stars, we plot a circle that refers to the case where the coefficients of the corresponding star are zero. Therefore, the radii of these circles are always  $\exp(0) = 1$ . If `scale=TRUE`, the stars are scaled so that they all have the same maximal ray length. In this case, the actual appearances of the circles differ, but they still refer to the no-effects case where all the coefficients are zero. Now the circles can be used to compare different stars based on their respective circles radii. The p-values beneath the covariate labels, which are given out if `test.rel=TRUE`, correspond to the distance between the circle and the star as a whole. They refer to a likelihood ratio test if all the coefficients from one covariate are zero (i.e. the variable is left out completely) and thus would lie exactly upon the circle.

The form of the circles can be modified by `col.circle`, `lwd.circle` and `lty.circle`.

By setting `globcircle=TRUE`, an additional circle can be drawn. The radii now correspond to a model, where the respective covariate is not included category-specific but globally. Therefore, the distance between this circle and the star as a whole corresponds to the p-value p-global that is given if `test.glob=TRUE`.

### Please note:

Regular fitting of cumulative logit models may fail because of the restrictions in the parameter space that have to be considered. If `partial=TRUE`, (sub)models with only one category-specific covariate, so-called partial proportional odds models, are fitted. Then at least estimates for every coefficient should be available. If `partial=TRUE`, the resulting effects of these (sub)models are plotted. It should be noted that in this case no coherent model is visualized. Also the p-values refer to the various submodels. For `partial=TRUE`, the p-values `p-rel` and `p-global` refer to tests of the corresponding partial proportional odds models against the proportional odds model.

It is strongly recommended to standardize metric covariates, display of effect stars can benefit greatly as in general differences between the coefficients are increased.

## Value

P-values are only available if the corresponding option is set TRUE.

<code>odds</code>	Odds or exponential coefficients of the cumulative logit model
<code>coefficients</code>	Coefficients of the cumulative logit model
<code>se</code>	Standard errors of the coefficients
<code>p_rel</code>	P-values of Likelihood-Ratio-Tests for the relevance of the explanatory covariates

p_global	P-values of Likelihood-Ratio-Tests whether the covariates need to be included category-specific
xlim	xlim values that were automatically produced. May be helpful if you want to specify your own xlim
ylim	ylim values that were automatically produced. May be helpful if you want to specify your own ylim

### Author(s)

Gunther Schauberger  
 <gunther@stat.uni-muenchen.de>  
<http://www.semsto.statistik.uni-muenchen.de/personen/doktoranden/schauberger/index.html>

### References

Tutz, G. and Schauberger, G. (2012): *Visualization of Categorical Response Models - from Data Glyphs to Parameter Glyphs*, Journal of Computational and Graphical Statistics 22(1), 156-177.

Gerhard Tutz (2012): *Regression for Categorical Data*, Cambridge University Press

### See Also

[star.sequential](#), [star.nominal](#)

### Examples

```
## Not run:
data(insolvency)

star.cumulative(Insolvency ~ Sector + Employees, insolvency, select = 2:4)

## End(Not run)
```

---

star.nominal

*Effect stars for multinomial logit models*

---

### Description

**The package EffectStars2 provides a more up-to-date implementation of effect stars!**

The function computes and visualizes multinomial logit models. The computation is done with help of the package **VGAM**. The visualization is based on the function [stars](#) from the package **graphics**.

**Usage**

```
star.nominal(formula, data, xij = NULL, conf.int = FALSE, symmetric = TRUE,
  pred.coding = "reference", printpvalues = TRUE, test.rel = TRUE, refLevel = 1,
  maxit = 100, scale = TRUE, nlines = NULL, select = NULL, catstar = TRUE,
  dist.x = 1, dist.y = 1, dist.cov = 1, dist.cat = 1, xpd = TRUE, main = "",
  lwd.stars = 1, col.fill = "gray90", col.circle = "black", lwd.circle = 1,
  lty.circle = "longdash", lty.conf = "dotted", cex.labels = 1, cex.cat = 0.8,
  xlim = NULL, ylim = NULL)
```

**Arguments**

formula	An object of class "formula". Formula for the multinomial logit model to be fitted and visualized.
data	An object of class "data.frame" containing the covariates used in formula.
xij	An object of class list, used if category-specific covariates are to be included. Every element is a formula referring to one of the category-specific covariates. For details see help for xij in <a href="#">vglm.control</a> and the details below.
conf.int	If TRUE, confidence intervals are drawn.
symmetric	Which side constraint for the coefficients in the multinomial logit model shall be used for the plot? Default TRUE uses symmetric side constraints, FALSE uses the reference category specified by refLevel. If category-specific covariates are specified using xij, automatically symmetric = FALSE is set. Symmetric side constraints are not possible in the case of category-specific covariates.
pred.coding	Which coding for categorical predictors with more than two categories is to be used? Default pred.coding="reference" uses the first category as reference category, the alternative pred.coding="effect" uses effect coding equivalent to symmetric side constraints. For pred.coding="effect" a star for every category is plotted, for pred.coding="reference" no star for the reference category is plotted.
printpvalues	If TRUE, p-values for the respective coefficients are printed besides the category labels. P-values are received by a Wald test.
test.rel	Provides a Likelihood-Ratio-Test to test the relevance of the explanatory covariates. The corresponding p-values will be printed behind the covariates labels. test.rel=FALSE might save a lot of time.
refLevel	Reference category for multinomial logit model. Ignored if symmetric=TRUE. See also <a href="#">multinomial</a> .
maxit	Maximal number of iterations to fit the multinomial logit model. See also <a href="#">vglm.control</a> .
scale	If TRUE, the stars are scaled to equal maximal ray length.
nlines	If specified, nlines gives the number of lines in which the effect stars are plotted.
select	Numeric vector to choose only a subset of the stars to be plotted. Default is to plot all stars. Numbers refer to total amount of predictors, including intercept and dummy variables.

catstar	A logical argument to specify if all category-specific effects in the model should be visualized with an additional star. Ignored if <code>xij=NULL</code> .
dist.x	Optional factor to increase/decrease distances between the centers of the stars on the x-axis. Values greater than 1 increase, values smaller than 1 decrease the distances.
dist.y	Optional factor to increase/decrease distances between the centers of the stars on the y-axis. Values greater than 1 increase, values smaller than 1 decrease the distances.
dist.cov	Optional factor to increase/decrease distances between the stars and the covariates labels above the stars. Values greater than 1 increase, values smaller than 1 decrease the distances.
dist.cat	Optional factor to increase/decrease distances between the stars and the category labels around the stars. Values greater than 1 increase, values smaller than 1 decrease the distances.
xpd	If FALSE, all plotting is clipped to the plot region, if TRUE, all plotting is clipped to the figure region, and if NA, all plotting is clipped to the device region. See also <a href="#">par</a> .
main	An overall title for the plot. See also <a href="#">plot</a> .
lwd.stars	Line width of the stars. See also <code>lwd</code> in <a href="#">par</a> .
col.fill	Color of background of the circle. See also <code>col</code> in <a href="#">par</a> .
col.circle	Color of margin of the circle. See also <code>col</code> in <a href="#">par</a> .
lwd.circle	Line width of the circle. See also <code>lwd</code> in <a href="#">par</a> .
lty.circle	Line type of the circle. See also <code>lty</code> in <a href="#">par</a> .
lty.conf	Line type of confidence intervals. Ignored, if <code>conf.int=FALSE</code> . See also <code>lty</code> in <a href="#">par</a> .
cex.labels	Size of labels for covariates placed above the corresponding star. See also <code>cex</code> in <a href="#">par</a> .
cex.cat	Size of labels for categories placed around the corresponding star. See also <code>cex</code> in <a href="#">par</a> .
xlim	Optional specification of the x coordinates ranges. See also <code>xlim</code> in <a href="#">plot.window</a>
ylim	Optional specification of the y coordinates ranges. See also <code>ylim</code> in <a href="#">plot.window</a>

## Details

The underlying models are fitted with the function `vglm` from the package **VGAM**. The family argument for `vglm` is `multinomial(parallel=FALSE)`.

The stars show the exponentials of the estimated coefficients. In multinomial logit models the exponential coefficients can be interpreted as odds. More precisely, for the model with symmetric side constraints, the exponential  $e^{\gamma_{rj}}$ ,  $r = 1, \dots, k$  represents the multiplicative effect of the covariate  $j$  on the odds  $\frac{P(Y=r|x)}{GM(x)}$  if  $x_j$  increases by one unit and  $GM(x)$  is the median response. For the model with reference category  $k$ , the exponential  $e^{\gamma_{rj}}$ ,  $r = 1, \dots, k-1$  represents the multiplicative effect of the covariate  $j$  on the odds  $\frac{P(Y=r|x)}{P(Y=k|x)}$  if  $x_j$  increases by one unit.



In addition to the stars, we plot a circle that refers to the case where the coefficients of the corresponding star are zero. Therefore, the radii of these circles are always  $\exp(0) = 1$ . If `scale=TRUE`, the stars are scaled so that they all have the same maximal ray length. In this case, the actual appearances of the circles differ, but they still refer to the no-effects case where all the coefficients are zero. Now the circles can be used to compare different stars based on their respective circles radii. The distances between the rays of a star and the circle correspond to the p-values that are printed beneath the category levels if `printpvalues=TRUE`. The closer a star ray lies to the no-effects circle, the more the p-value is increased.

The p-values beneath the covariate labels, which are given if `test.rel=TRUE`, correspond to the distance between the circle and the star as a whole. They refer to a likelihood ratio test if all the coefficients from one covariate are zero (i.e. the variable is left out completely) and thus would lie exactly upon the circle.

The appearance of the circles can be modified by `col.circle`, `lwd.circle` and `lty.circle`.

The argument `xij` is important because it has to be used to include category-specific covariates. If its default `xij=NULL` is kept, an ordinary multinomial logit model without category-specific covariates is fitted. If category-specific covariates are to be included, attention has to be paid to the exact usage of `xij`. Our `xij` argument is identical to the `xij` argument used in the embedded `vglm` function. For details see also `vglm.control`. The data are thought to be present in a wide format, i.e. a category-specific covariate consists of `k` columns. Before calling `star.nominal`, the values for the reference category (defined by `refLevel`) have to be subtracted from the values of the further categories. Additionally, the resulting variable for the first response category (but not the reference category) has to be duplicated. This duplicate should be denoted by an appropriate name for the category-specific variable, independent from the different response categories. It will be used as an assignment variable for the corresponding coefficient of the covariate and has to be included in to the formula. For every category-specific covariate, a formula has to be specified in the `xij` argument. On the left hand side of that formula, the assignment variable has to be placed. On the right hand side, the variables containing the differences from the values for the reference category are written. So the left hand side of the formula contains `k-1` terms. The order of these terms has to be chosen according to the order of the response categories, ignoring the reference category. Examples for effect stars for models with category-specific covariates are received by typing `vignette("election")` or `vignette("plebiscite")`.

It is strongly recommended to standardize metric covariates, display of effect stars can benefit greatly as in general differences between the coefficients are increased.

## Value

P-values are only available if the corresponding option is set `TRUE`.

`catspec` and `catspecse` are only available if `xij` is specified.

<code>odds</code>	Odds or exponential coefficients of the multinomial logit model
<code>coefficients</code>	Coefficients of the multinomial logit model
<code>se</code>	Standard errors of the coefficients
<code>pvalues</code>	P-values of Wald tests for the respective coefficients
<code>catspec</code>	Coefficients for the category-specific covariates

catspecse	Standard errors for the coefficients for the category-specific covariates
p_rel	P-values of Likelihood-Ratio-Tests for the relevance of the explanatory covariates
xlim	xlim values that were automatically produced. May be helpful if you want to specify your own xlim
ylim	ylim values that were automatically produced. May be helpful if you want to specify your own ylim

**Author(s)**

Gunther Schaubberger  
 <gunther@stat.uni-muenchen.de>  
<http://www.semsto.statistik.uni-muenchen.de/personen/doktoranden/schaubberger/index.html>

**References**

Tutz, G. and Schaubberger, G. (2012): *Visualization of Categorical Response Models - from Data Glyphs to Parameter Glyphs*, Journal of Computational and Graphical Statistics 22(1), 156-177.

Gerhard Tutz (2012): *Regression for Categorical Data*, Cambridge University Press

**See Also**

[star.sequential](#), [star.cumulative](#)

**Examples**

```
## Not run:
data(election)

# simple multinomial logit model
star.nominal(Partychoice ~ Age + Religion + Democracy + Pol.Interest +
             Unemployment + Highschool + Union + West + Gender, election)

# Use effect coding for the categorical predictor religion
star.nominal(Partychoice ~ Age + Religion + Democracy + Pol.Interest +
             Unemployment + Highschool + Union + West + Gender, election,
             pred.coding = "effect")

# Use reference category "FDP" instead of symmetric side constraints
star.nominal(Partychoice ~ Age + Religion + Democracy + Pol.Interest +
             Unemployment + Highschool + Union + West + Gender, election,
             refLevel = 3, symmetric = FALSE)

# Use category-specific covariates, subtract values for reference
# category CDU
election[,13:16] <- election[,13:16] - election[,12]
election[,18:21] <- election[,18:21] - election[,17]
election[,23:26] <- election[,23:26] - election[,22]
```

```

election[,28:31] <- election[,28:31] - election[,27]

election$Social <- election$Social_SPD
election$Immigration <- election$Immigration_SPD
election$Nuclear <- election$Nuclear_SPD
election$Left_Right <- election$Left_Right_SPD

star.nominal(Partychoice ~ Social + Immigration + Nuclear + Left_Right + Age +
Religion + Democracy + Pol.Interest + Unemployment + Highschool + Union + West +
Gender, data = election,
xij = list(Social ~ Social_SPD + Social_FDP + Social_Greens + Social_Left,
Immigration ~ Immigration_SPD + Immigration_FDP + Immigration_Greens + Immigration_Left,
Nuclear ~ Nuclear_SPD + Nuclear_FDP + Nuclear_Greens + Nuclear_Left,
Left_Right ~ Left_Right_SPD + Left_Right_FDP + Left_Right_Greens + Left_Right_Left),
symmetric = FALSE)

## End(Not run)

```

---

star.sequential	<i>Effect stars for sequential logit models</i>
-----------------	-------------------------------------------------

---

## Description

**The package EffectStars2 provides a more up-to-date implementation of effect stars!**

The function computes and visualizes sequential logit models. The computation is done with help of the package **VGAM**. The visualization is based on the function [stars](#) from the package **graphics**.

## Usage

```

star.sequential(formula, data, global = NULL, test.rel = TRUE, test.glob = FALSE,
globcircle = FALSE, maxit = 100, scale = TRUE, nlines = NULL, select = NULL,
dist.x = 1, dist.y = 1, dist.cov = 1, dist.cat = 1, xpd = TRUE, main = "",
col.fill = "gray90", col.circle = "black", lwd.circle = 1,
lty.circle = "longdash", col.global = "black", lwd.global = 1,
lty.global = "dotdash", cex.labels = 1, cex.cat = 0.8, xlim = NULL,
ylim = NULL)

```

## Arguments

formula	An object of class “formula”. Formula for the sequential logit model to be fitted and visualized.
data	An object of class “data.frame” containing the covariates used in formula.
global	Numeric vector to choose a subset of predictors to be included with global coefficients. Default is to include all coefficients category-specific. Numbers refer to total amount of predictors, including intercept and dummy variables.
test.rel	Provides a Likelihood-Ratio-Test to test the relevance of the explanatory covariates. The corresponding p-values will be printed as p-rel. test.rel=FALSE might save a lot of time.

<code>test.glob</code>	Provides a Likelihood-Ratio-Test to test if a covariate has to be included as a category-specific covariate (in contrast to being global). The corresponding p-values will be printed as <code>p-global</code> . <code>test.glob=FALSE</code> and <code>globcircle=FALSE</code> might save a lot of time.
<code>globcircle</code>	If TRUE, additional circles that represent the global effects of the covariates are plotted. <code>test.glob=FALSE</code> and <code>globcircle=FALSE</code> might save a lot of time.
<code>maxit</code>	Maximal number of iterations to fit the sequential logit model. See also <a href="#">vglm.control</a> .
<code>scale</code>	If TRUE, the stars are scaled to equal maximal ray length.
<code>nlines</code>	If specified, <code>nlines</code> gives the number of lines in which the effect stars are plotted.
<code>select</code>	Numeric vector to choose only a subset of the stars to be plotted. Default is to plot all stars. Numbers refer to total amount of predictors, including intercept and dummy variables.
<code>dist.x</code>	Optional factor to increase/decrease distances between the centers of the stars on the x-axis. Values greater than 1 increase, values smaller than 1 decrease the distances.
<code>dist.y</code>	Optional factor to increase/decrease distances between the centers of the stars on the y-axis. Values greater than 1 increase, values smaller than 1 decrease the distances.
<code>dist.cov</code>	Optional factor to increase/decrease distances between the stars and the covariates labels above the stars. Values greater than 1 increase, values smaller than 1 decrease the distances.
<code>dist.cat</code>	Optional factor to increase/decrease distances between the stars and the category labels around the stars. Values greater than 1 increase, values smaller than 1 decrease the distances.
<code>xpd</code>	If FALSE, all plotting is clipped to the plot region, if TRUE, all plotting is clipped to the figure region, and if NA, all plotting is clipped to the device region. See also <a href="#">par</a> .
<code>main</code>	An overall title for the plot. See also <a href="#">plot</a> .
<code>col.fill</code>	Color of background of the circle. See also <code>col</code> in <a href="#">par</a> .
<code>col.circle</code>	Color of margin of the circle. See also <code>col</code> in <a href="#">par</a> .
<code>lwd.circle</code>	Line width of the circle. See also <code>lwd</code> in <a href="#">par</a> .
<code>lty.circle</code>	Line type of the circle. See also <code>lty</code> in <a href="#">par</a> .
<code>col.global</code>	Color of margin of the global effects circle. See also <code>col</code> in <a href="#">par</a> . Ignored, if <code>globcircle = FALSE</code> .
<code>lwd.global</code>	Line width of the global effects circle. See also <code>lwd</code> in <a href="#">par</a> . Ignored, if <code>globcircle = FALSE</code> .
<code>lty.global</code>	Line type of the global effects circle. See also <code>lty</code> in <a href="#">par</a> . Ignored, if <code>globcircle = FALSE</code> .
<code>cex.labels</code>	Size of labels for covariates placed above the corresponding star. See also <code>cex</code> in <a href="#">par</a> .
<code>cex.cat</code>	Size of labels for categories placed around the corresponding star. See also <code>cex</code> in <a href="#">par</a> .
<code>xlim</code>	Optional specification of the x coordinates ranges. See also <code>xlim</code> in <a href="#">plot.window</a>
<code>ylim</code>	Optional specification of the y coordinates ranges. See also <code>ylim</code> in <a href="#">plot.window</a>

## Details

The underlying models are fitted with the function `vglm` from the package **VGAM**. The family argument for `vglm` is `sratio(parallel=FALSE)`.

The stars show the exponentials of the estimated coefficients. In sequential logit models the exponential coefficients can be interpreted as odds. More precisely, the exponential  $e^{\gamma_{rj}}$ ,  $r = 1, \dots, k-1$  represents the multiplicative effect of the covariate  $j$  on the continuation ratio odds  $\frac{P(Y=r|x)}{P(Y>r|x)}$  if  $x_j$  increases by one unit.

In addition to the stars, we plot a circle that refers to the case where the coefficients of the corresponding star are zero. Therefore, the radii of these circles are always  $exp(0) = 1$ . If `scale=TRUE`, the stars are scaled so that they all have the same maximal ray length. In this case, the actual appearances of the circles differ, but they still refer to the no-effects case where all the coefficients are zero. Now the circles can be used to compare different stars based on their respective circles radii. The p-values beneath the covariate labels, which are given out if `test.rel=TRUE`, correspond to the distance between the circle and the star as a whole. They refer to a likelihood ratio test if all the coefficients from one covariate are zero (i.e. the variable is left out completely) and thus would lie exactly upon the circle.

The appearance of the circles can be modified by `col.circle`, `lwd.circle` and `lty.circle`.

By setting `globcircle=TRUE`, an additional circle can be drawn. The radii now correspond to a model, where the respective covariate is not included category-specific but globally. Therefore, the distance between this circle and the star as a whole corresponds to the p-value p-global that is given if `test.glob=TRUE`.

It is strongly recommended to standardize metric covariates, display of effect stars can benefit greatly as in general differences between the coefficients are increased.

## Value

P-values are only available if the corresponding option is set TRUE.

<code>odds</code>	Odds or exponential coefficients of the sequential logit model
<code>coefficients</code>	Coefficients of the sequential logit model
<code>se</code>	Standard errors of the coefficients
<code>p_rel</code>	P-values of Likelihood-Ratio-Tests for the relevance of the explanatory covariates
<code>p_global</code>	P-values of Likelihood-Ratio-Tests whether the covariates need to be included category-specific
<code>xlim</code>	<code>xlim</code> values that were automatically produced. May be helpful if you want to specify your own <code>xlim</code>
<code>ylim</code>	<code>ylim</code> values that were automatically produced. May be helpful if you want to specify your own <code>ylim</code>

**Author(s)**

Gunther Schaubberger  
 <gunther@stat.uni-muenchen.de>  
<http://www.semsto.statistik.uni-muenchen.de/personen/doktoranden/schauberger/index.html>

**References**

Tutz, G. and Schaubberger, G. (2012): *Visualization of Categorical Response Models - from Data Glyphs to Parameter Glyphs*, Journal of Computational and Graphical Statistics 22(1), 156-177.

Gerhard Tutz (2012): *Regression for Categorical Data*, Cambridge University Press

**See Also**

[star.nominal](#), [star.cumulative](#)

**Examples**

```
## Not run:
data(insolvency)

star.sequential(Insolvency ~ Sector + Legal + Pecuniary_Reward + Seed_Capital
+ Debt_Capital + Employees, insolvency, test.glob = FALSE, globcircle = TRUE, dist.x = 1.3)

## End(Not run)
```

---

womenlabour

*Canadian Women's Labour-Force Participation*

---

**Description**

The data are from a 1977 survey of the Canadian population.

**Usage**

```
data(womenlabour)
```

**Format**

A data frame with 263 observations on the following 4 variables.

Participation Labour force participation with levels fulltime, not.work and parttime

IncomeHusband Husband's income in 1000 \$

Children Presence of children in household with levels absent and present

Region Region with levels Atlantic, BC, Ontario, Prairie and Quebec

## Source

R package car: [Womenlf](#)

## References

*Social Change in Canada Project*. York Institute for Social Research.

Fox, J. (2008): *Applied Regression Analysis and Generalized Linear Models*, Second Edition.

## Examples

```
## Not run:  
data(womenlabour)  
womenlabour$IncomeHusband <- scale(womenlabour$IncomeHusband)  
  
star.nominal(Participation ~ IncomeHusband + Children + Region, womenlabour)  
  
## End(Not run)
```

# Index

- \*Topic **categorical data**
  - EffectStars, 5
- \*Topic **cumulative logit model**
  - EffectStars, 5
  - star.cumulative, 11
- \*Topic **datasets**
  - alligator, 2
  - BEPS, 3
  - coffee, 4
  - election, 5
  - insolvency, 8
  - PID, 9
  - plebiscite, 10
  - womenlabour, 22
- \*Topic **multinomial logit model**
  - EffectStars, 5
  - star.nominal, 14
- \*Topic **multinomial response**
  - alligator, 2
  - BEPS, 3
  - coffee, 4
  - EffectStars, 5
  - election, 5
  - PID, 9
  - plebiscite, 10
  - star.nominal, 14
  - womenlabour, 22
- \*Topic **ordinal response**
  - EffectStars, 5
  - insolvency, 8
  - star.cumulative, 11
  - star.sequential, 19
- \*Topic **package**
  - EffectStars, 5
- \*Topic **sequential logit model**
  - EffectStars, 5
  - star.sequential, 19
- \*Topic **star plot**
  - star.cumulative, 11
  - star.nominal, 14
  - star.sequential, 19
- alligator, 2
- BEPS, 3, 3
- Chile, 10
- coffee, 4
- EffectStars, 5
- EffectStars-package (EffectStars), 5
- election, 5
- insolvency, 8
- multinomial, 15
- nes96, 9
- par, 12, 16, 20
- PID, 9
- plebiscite, 10
- plot, 12, 16, 20
- plot.window, 12, 16, 20
- star.cumulative, 5, 11, 18, 22
- star.nominal, 5, 14, 14, 22
- star.sequential, 5, 14, 18, 19
- stars, 11, 14, 19
- vglm, 13, 16, 21
- vglm.control, 12, 15, 17, 20
- womenlabour, 22
- Womenlf, 23