Package 'DGLMExtPois'

May 3, 2020

Type Package
Title Double Generalized Linear Models Extending Poisson Regression
Version 0.1.3
Maintainer Francisco Martinez <fmartin@ujaen.es></fmartin@ujaen.es>
Description Model estimation, dispersion testing and diagnosis of hyper-Poisson Saez-Castillo, A.J. and Conde-Sanchez, A. (2013) <doi:10.1016 j.csda.2012.12.009=""> and Conway-Maxwell-Poisson Huang, A. (2017) <doi:10.1177 1471082x17697749=""> regression models.</doi:10.1177></doi:10.1016>
Depends R (>= $3.5.0$)
License GPL-2
Encoding UTF-8
LazyData true
Imports nloptr (>= 1.2.1), COMPoissonReg, progress
RoxygenNote 7.1.0
NeedsCompilation no
Author Antonio Jose Saez-Castillo [aut], Antonio Conde-Sanchez [aut], Francisco Martinez [aut, cre]
Repository CRAN
Date/Publication 2020-05-03 12:40:02 UTC
AIC 2 AIC_CMP 3 Bids 3 confint.glm_CMP 4 confint.glm_hP 5 DGLMExtPois 6 expected 6 glm.CMP 7

2 AIC

hP						 												12
lrtest						 												13
plots						 												14
predict.glm_CMP .						 												15
predict.glm_hP						 												16
residuals						 												17
summary.glm_CMP	٠.					 												19
summary.glm_hP .						 												19
																		21

Index

AIC

AIC and BIC for hyper-Poisson Fits

Description

Computes the Akaike's information criterion or the Shwarz's Bayesian criterion for hyper-Poisson Fits

Usage

```
## S3 method for class 'glm_hP'
AIC(object, ..., k = 2)
## S3 method for class 'glm_hP'
BIC(object, ...)
```

Arguments

object an object of class "glm_hP", typically the result of a call to glm. hP.
optionally more fitted model objects.
numeric, the *penalty* per parameter to be used; the default k = 2 is the classical AIC.

AIC_CMP 3

AIC_CMP

AIC and BIC for COM-Poisson Fitted Models

Description

Computes the Akaike's information criterion or the Shwarz's Bayesian criterion for COM-Poisson Fits

Usage

```
## S3 method for class 'glm_CMP'
AIC(object, ..., k = 2)
## S3 method for class 'glm_CMP'
BIC(object, ...)
```

Arguments

object an object of class "glm_CMP", typically the result of a call to glm. CMP.
 optionally more fitted model objects.
 numeric, the *penalty* per parameter to be used; the default k = 2 is the classical AIC.

Examples

Bids

Bids Received by U.S. Firms

Description

A dataset with bids received by U.S. firms.

Usage

Bids

4 confint.glm_CMP

Format

A data frame with 126 rows and 13 variables:

docno doc no.weeks weeksnumbids count

takeover delta(1 if taken over)

bidprem bid Premium

insthold institutional holdings
size size measured in billions
leglrest legal restructuring
rearest real restructuring
finrest financial restructuring
regulatn regulation

whtknght white knight

Source

Jaggia, Sanjiv and Satish Thosar (1993) "Multiple Bids as a Consequence of Target Management Resistance", Review of Quantitative Finance and Accounting, 447–457.

Cameron, A.C. and Per Johansson (1997) "Count Data Regression Models using Series Expansions: with Applications", Journal of Applied Econometrics, 12, may, 203–223.

confint.glm_CMP

Confidence Intervals for glm_hP Fits

Description

Computes confidence intervals for one or more parameters in a glm_CMP object.

Usage

```
## S3 method for class 'glm_CMP'
confint(object, parm, level = 0.95, ...)
```

Arguments

object a fitted object of class inheriting from "glm_CMP".

parm a specification of which parameters are to be given confidence intervals, either

a vector of numbers or a vector of names. If missing, all parameters are consid-

ered.

level the confidence level required.

additional argument(s) for methods.

confint.glm_hP 5

Value

A matrix (or vector) with columns giving lower and upper confidence limits for each beta parameter. These will be labelled as (1-level)/2 and 1 - (1-level)/2 in (by default 2.5% and 97.5%).

Examples

confint.glm_hP

Confidence Intervals for glm_hP Fits

Description

Computes confidence intervals for one or more parameters in a "glm_hP" object.

Usage

```
## S3 method for class 'glm_hP'
confint(object, parm, level = 0.95, ...)
```

Arguments

object a fitted object of class inheriting from "glm_hP".

parm a specification of which parameters are to be given confidence intervals, either

a vector of numbers or a vector of names. If missing, all parameters are consid-

ered.

level the confidence level required.

... additional argument(s) for methods.

Value

A matrix (or vector) with columns giving lower and upper confidence limits for each beta parameter. These will be labelled as (1-level)/2 and 1 - (1-level)/2 in % (by default 2.5% and 97.5%).

6 expected

DGLMExtPois	DGLMExtPois: Double Generalized Linear Models Extending Pois-
	son Regression.

Description

The DGLMExtPois package provides two categories of important functions: glm.hP and glm.CMP.

expected	Expected Probabilities and Frequencies for the hyper-Poisson and COM-Poisson Model

Description

The hP_expected and CMP_expected functions calculate the probability distribution of the count response variable Y for each observation and obtain the corresponding expected frequencies. It is an informal way to evaluate the fit of the hP or CMP model by comparing the predicted distribution of counts with the observed distribution.

Usage

```
hP_expected(object)
CMP_expected(object)
```

Arguments

object a fitted object of class inheriting from "glm_hP" or "glm_CMP".

Details

The average expected probabilities are computed as

$$\overline{(Pr)}(y=k) = \frac{1}{n} \sum_{i=1}^{n} \widehat{Pr}(y_i = k|x_i)$$

The expected frequencies are obtained by multiplying by n.

Two measures are offered for summarizing the comparison between expected and observed frequencies: the sum of the absolute value of differences and the sum of the square of differences (similar to the Pearson statistic of goodness of fit).

glm.CMP 7

Value

A list containing the following components:

frequencies the expected counts for the hP or CMP fit.

observed_freq the observed distribution.

probabilities the expected distribution for the hP or CMP fit.

dif sum of the absolute value of differences between frequencies and observed_freq.

chi2 sum of the square of differences between frequencies and observed_freq.

References

Hilbe, J. M. (2011). Negative Binomial Regression. (2nd ed.). Cambridge University Press.#' Long, J. S. & Freese, J. (2014). Regression Models for Categorical Dependent Variables using STATA. (3rd ed.). Stata Press.

Examples

glm.CMP

Fit a COM-Poisson Double Generalized Linear Model

Description

The glm. CMP function is used to fit a COM-Poisson double generalized linear model with a log-link for the mean (mu) and the dispersion parameter (nu).

Usage

```
glm.CMP(
  formula.mu,
  formula.nu,
  init.beta = NULL,
  init.delta = NULL,
  data,
  weights,
  subset,
```

8 glm.CMP

```
na.action,
maxiter_series = 1000,
tol = 0,
offset,
opts = NULL,
model.mu = TRUE,
model.nu = TRUE,
x = FALSE,
y = TRUE,
z = FALSE
```

Arguments

formula.mu regression formula linked to log(mu) regression formula linked to log(nu)

init.beta initial values for regression coefficients of beta.init.delta initial values for regression coefficients of delta.

data an optional data frame, list or environment (or object coercible by as. data. frame

to a data frame) containing the variables in the model. If not found in data, the variables are taken from environment (formula), typically the environment

from which glm. CMP is called.

weights an optional vector of 'prior weights' to be used in the fitting process. Should be

NULL or a numeric vector.

subset an optional vector specifying a subset of observations to be used in the fitting

process.

na.action a function which indicates what should happen when the data contain NAs. The

default is set by the na.action setting of options, and is na.fail if that is unset. The 'factory-fresh' default is na.omit. Another possible value is NULL,

no action. Value na. exclude can be useful.

maxiter_series Maximum number of iterations to perform in the calculation of the normalizing

constant.

tol tolerance with default zero meaning to iterate until additional terms to not change

the partial sum in the calculation of the normalizing constant.

offset this can be used to specify an a priori known component to be included in the

linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases. One or more offset terms can be included in the formula instead or as well, and if more than one is specified their sum is used.

 $See \ {\tt model.offset}.$

opts a list with options to the optimizer, nloptr, that fits the model. See, the opts

parameter of nloptr for further details.

model.mu a logical value indicating whether the mu model frame should be included as a

component of the returned value.

model.nu a logical value indicating whether the *nu model frame* should be included as a

component of the returned value.

glm.CMP

х	logical value indicating whether the mu model matrix used in the fitting process should be returned as a component of the returned value.
у	logical value indicating whether the response vector used in the fitting process should be returned as a component of the returned value.
z	logical value indicating whether the nu model matrix used in the fitting process should be returned as a component of the returned value.

Details

Fit a COM-Poisson double generalized linear model using as optimizer the NLOPT_LD_SLSQP algorithm of function nloptr.

Value

glm.CMP returns an object of class "glm_CMP". The function summary can be used to obtain or print a summary of the results. An object of class "glm_CMP" is a list containing at least the following components:

```
coefficients
                   a named vector of coefficients.
residuals
                   the residuals, that is response minus fitted values.
fitted.values
                  the fitted mean values.
linear.predictors
                   the linear fit on link scale.
call
                   the matched call.
offset
                   the offset vector used.
weights
                   the weights initially supplied, a vector of 1s if none were.
                  if requested (the default) the y vector used.
matrix.mu
                  if requested, the mu model matrix.
                  if requested, the nu model matrix.
```

References

matrix.nu
model.mu

model.nu

Huang, A. (2017). Mean-parametrized Conway–Maxwell–Poisson regression models for dispersed counts. Statistical Modelling 17, 359–380.

Johnson, S. G. (2018). The nlopt nonlinear-optimization package

Examples

if requested (the default) the mu model frame.

if requested (the default) the nu model frame.

10 glm.hP

glm.hP

Fit a hyper-Poisson Double Generalized Linear Model

Description

The glm.hP function is used to fit a hyper-Poisson double generalized linear model with a log-link for the mean (mu) and the dispersion parameter (gamma).

Usage

```
glm.hP(
  formula.mu,
  formula.gamma,
  init.beta = NULL,
  init.delta = NULL,
  data,
 weights,
  subset,
  na.action,
 maxiter_series = 1000,
  tol = 0,
  offset,
 opts = NULL,
 model.mu = TRUE,
 model.gamma = TRUE,
 x = FALSE,
 y = TRUE,
 z = FALSE
)
```

Arguments

formula.mu	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted.
formula.gamma	regression formula linked to log(gamma)

init.beta initial values for regression coefficients of beta.
init.delta initial values for regression coefficients of delta.

an optional data frame, list or environment (or object that can be coerced by as.data.frame to a data frame) containing the variables in the model. If not found in data, the variables are taken from environment(formula), typically

the environment from which glm. hP is called.

weights an optional vector of 'prior weights' to be used in the fitting process. Should be

NULL or a numeric vector.

subset an optional vector specifying a subset of observations to be used in the fitting

process.

glm.hP

na.action	a function which indicates what should happen when the data contain NAs. The default is set by the na.action setting of options, and is na.fail if that is unset. The 'factory-fresh' default is na.omit. Another possible value is NULL, no action. Value na.exclude can be useful.
maxiter_series	Maximum number of iterations to perform in the calculation of the normalizing constant.
tol	tolerance with default zero meaning to iterate until additional terms to not change the partial sum in the calculation of the normalizing constant.
offset	this can be used to specify an a priori known component to be included in the linear predictor during fitting. This should be NULL or a numeric vector of length equal to the number of cases. One or more offset terms can be included in the formula instead or as well, and if more than one is specified their sum is used. See model.offset.
opts	a list with options to the optimizer, nloptr, that fits the model. See, the opts parameter of nloptr for further details.
model.mu	a logical value indicating whether the mu $model$ $frame$ should be included as a component of the returned value.
model.gamma	a logical value indicating whether the <i>gamma model frame</i> should be included as a component of the returned value.
x	logical value indicating whether the mu model matrix used in the fitting process should be returned as a component of the returned value.
у	logical value indicating whether the response vector used in the fitting process should be returned as a component of the returned value.
z	logical value indicating whether the gamma model matrix used in the fitting process should be returned as a component of the returned value.

Details

Fit a hyper-Poisson double generalized linear model using as optimizer the NLOPT_LD_SLSQP algorithm of function nloptr.

Value

glm.hP returns an object of class "glm_hP". The function summary can be used to obtain or print a summary of the results.

The generic accessor functions coef, fitted.values and residuals can be used to extract various useful features of the value returned by glm. hP.

 $weights\ extracts\ a\ vector\ of\ weights, one\ for\ each\ case\ in\ the\ fit\ (after\ subsetting\ and\ na.\ action).$

An object of class "glm_hP" is a list containing at least the following components:

coefficients a named vector of coefficients.

residuals the residuals, that is response minus fitted values.

fitted.values the fitted mean values.

linear.predictors

the linear fit on link scale.

12 hP

call the matched call. offset the offset vector used. weights the weights initially supplied, a vector of 1s if none were. df.residual the residual degrees of freedom. df.null the residual degrees of freedom for the null model. if requested (the default) the y vector used. У matrix.mu if requested, the mu model matrix. matrix.gamma if requested, the gamma model matrix. if requested (the default) the mu model frame. model.mu

References

model.gamma

Saez-Castillo, A.J. & Conde-Sanchez, A. (2013). "A hyper-Poisson regression model for overdispersed and underdispersed count data", Computational Statistics & Data Analysis, vol. 61(C), pages 148–157

if requested (the default) the gamma model frame.

Johnson, S. G. (2018). The nlopt nonlinear-optimization package

Examples

hΡ

The hyper-Poisson Distribution

Description

Density, distribution function and random generation for the hyper-Poisson distribution with parameters gamma and lambda.

Usage

```
dhP(x, gamma, lambda)
phP(q, gamma, lambda, lower.tail = TRUE)
rhP(n, gamma, lambda)
```

lrtest 13

Arguments

X	vector of (non-negative integer) quantiles.
gamma	dispersion parameter. Must be strictly positive.
lambda	location parameter. Must be strictly positive.
q	vector of quantiles.
lower.tail	logical; if TRUE (default), probabilities are $P[X \leq x]$, otherwise, $P[X > x]$.
n	number of random values to return.

Value

dhP gives the density, phP gives the distribution function and rhP generates random deviates.

Invalid gamma or lambda will result in return value NaN, with a warning.

The length of the result is determined by n for rhP, and is the maximum of the lengths of the numerical arguments for the other functions.

The numerical arguments other than n are recycled to the length of the result. Only the first element of the logical arguments is used.

Examples

```
dhP(3, 15, 2)
phP(3, 15, 2)
rhP(10, 15, 2)
```

lrtest

Likelihood Ratio Test for Nested glm_CMP and glm_hP Fits

Description

Performs the likelihood ratio chi-squared test to compare nested models.

Usage

```
lrtest(object1, object2)
```

Arguments

```
object1, object2
```

fitted objects of classes inheriting from "glm_CMP" or "glm_hP"

Details

The test statistics is calculated as $2(llik - llik_0)$. The test statistics has a chi-squared distribution with r degrees of freedom, where r is the difference in the number of parameters between the full and null models.

14 plots

Value

A list with class "lrtest" containing the following components:

```
statistics the value of the statistic.

df the degrees of freedom.
p-value the p-value for the test.
```

Examples

plots

Plot Diagnostics for glm_hP and glm_CMP Objects

Description

Two plots are currently available: a plot of residuals against fitted values and a Normal Q-Q plot.

Usage

```
## S3 method for class 'glm_hP'
plot(
    X,
    type = c("quantile", "pearson", "response"),
    ask = prod(graphics::par("mfcol")) < 2 && grDevices::dev.interactive(),
    ...
)

## S3 method for class 'glm_CMP'
plot(
    X,
    type = c("quantile", "pearson", "response"),
    ask = prod(graphics::par("mfcol")) < 2 && grDevices::dev.interactive(),
    ...
)</pre>
```

predict.glm_CMP 15

Arguments

```
x glm_hP or glm_CMP object, typically the result of glm.hP or glm.CMP.

type the type of residuals. The alternatives are: "quantile" (default), "pearson" and "response". Can be abbreviated.

ask logical; if TRUE, the user is asked before each plot, see par(ask=.).

other parameters to be passed through to plotting functions.
```

Examples

predict.glm_CMP

Predict Method for glm_CMP Fits

Description

Obtains predictions from a fitted glm_CMP object.

Usage

```
## S3 method for class 'glm_CMP'
predict(object, newdata = NULL, type = c("link", "response"), ...)
```

Arguments

object	a fitted object of class inheriting from "glm_CMP".
newdata	optionally, a data frame in which to look for variables with which to predict. If omitted, the fitted linear predictors are used.
type	the type of prediction required. The default is on the scale of the linear predictors; the alternative "response" is on the scale of the response variable.
	further arguments passed to or from other methods.

16 predict.glm_hP

Value

A vector with the prediction means.

Examples

predict.glm_hP

Predict Method for glm_hP Fits

Description

Obtains predictions from a fitted "glm_hP" object.

Usage

```
## S3 method for class 'glm_hP'
predict(object, newdata = NULL, type = c("link", "response"), ...)
```

Arguments

object a fitted object of class inheriting from "glm_hP".

newdata optionally, a data frame in which to look for variables with which to predict. If omitted, the fitted linear predictors are used.

type the type of prediction required. The default is on the scale of the linear predictors; the alternative "response" is on the scale of the response variable.

further arguments passed to or from other methods.

Value

A vector with the prediction means.

residuals 17

residuals	Extract and Visualize hyper-Poisson and COM-Poisson Model Residuals

Description

residuals is a method which extracts model residuals from a "glm_hP" or "glm_CMP" object, commonly returned by glm.hP or glm.CMP. Optionally, it produces a half normal plot with a simulated envelope of the residuals.

Usage

```
## S3 method for class 'glm_hP'
residuals(
  object,
  type = c("pearson", "response", "quantile"),
  envelope = FALSE,
  rep = 19,
  title = "Simulated Envelope of Residuals",
)
## S3 method for class 'glm_CMP'
residuals(
 object,
  type = c("pearson", "response", "quantile"),
  envelope = FALSE,
 rep = 19,
  title = "Simulated Envelope of Residuals",
)
```

Arguments

object	an object of class "glm_hP" or "glm_CMP", typically the result of a call to glm.hP or glm.CMP.
type	the type of residuals which should be returned. The alternatives are: "pearson" (default), "response" and "quantile". Can be abbreviated.
envelope	a logical value indicating whether the envelope should be computed.
rep	number of replications for envelope construction. Default is 19, that is the smallest 95 percent band that can be built.
title	a string indicating the main title of the envelope.
	further arguments passed to or from other methods.

18 residuals

Details

The response residuals $(r_i = y_i - \mu_i)$, Pearson residuals $(r_i^P = r_i/\sigma_i)$ or randomized quantile residuals are computed. The randomized quantile residuals are obtained computing the cumulative probabilities that the fitted model being less than y and less or equal than y. A random value from a uniform distribution between both probabilities is generated and the value of the residual is the standard normal variate with the same cumulative probability. Four replications of the quantile residuals are recommended because of the random component (see Dunn and Smyth, 1996 for more details).

The functions plot.glm_hP and plot.glm_CMP generate a residuals against fitted values plot and a Normal Q-Q plot.

The Normal Q-Q plot may show an unsatisfactory pattern of the Pearson residuals of a fitted model: then we are led to think that the model is incorrectly specified. The half normal plot with simulated envelope indicates that, under the distribution of the response variable, the model is fine when only a few points fall off the envelope.

Value

Residual values.

References

Dunn, P.K., Smyth, G.K. (1996). Randomized quantile residuals. Journal of Computational and Graphical Statistics, 5(3), 236-244.

Atkinson, A. (1981). Two graphical displays for outlying and influential observations in regression. Biometrika, 68 (1), 13–20.

See Also

plots

summary.glm_CMP 19

summary.glm_CMP

Summarizing COM-Poisson Fits

Description

These functions are all methods for class "glm_CMP" or summary.glm_CMP objects.

Usage

```
## S3 method for class 'glm_CMP'
summary(object, ...)

## S3 method for class 'summary.glm_CMP'
print(
    x,
    digits = max(3, getOption("digits") - 3),
    signif.stars = getOption("show.signif.stars"),
    ...
)
```

Arguments

```
    an object of class "glm_CMP", usually, a result of a call to glm. CMP.
    further arguments passed to or from other methods.
    an object of class "summary.glm_CMP", usually, a result of a call to summary.glm_CMP.
    the number of significant digits to use when printing.
    logical. If TRUE, 'significance stars' are printed for each coefficient.
```

Examples

summary.glm_hP

Summarizing hyper-Poisson Fits

Description

These functions are all methods for class "glm_hP" or summary.glm_hP objects.

20 summary.glm_hP

Usage

```
## $3 method for class 'glm_hP'
summary(object, ...)

## $3 method for class 'summary.glm_hP'
print(
    x,
    digits = max(3, getOption("digits") - 3),
    signif.stars = getOption("show.signif.stars"),
    ...
)
```

Arguments

```
object an object of class "glm_hP", usually, a result of a call to glm.hP.

further arguments passed to or from other methods.

x an object of class "summary.glm_hP", usually, a result of a call to summary.glm_hP.

digits the number of significant digits to use when printing.

signif.stars logical. If TRUE, 'significance stars' are printed for each coefficient.
```

Index

```
*Topic datasets
    Bids, 3
AIC, 2
AIC.glm_CMP (AIC_CMP), 3
AIC_CMP, 3
as.data.frame, 8, 10
BIC.glm_CMP (AIC_CMP), 3
BIC.glm_hP(AIC), 2
Bids, 3
CMP_expected (expected), 6
coef, 11
confint.glm_CMP, 4
confint.glm_hP, 5
DGLMExtPois, 6
dhP (hP), 12
expected, 6
fitted.values, 11
glm.CMP, 3, 7, 15, 17
glm. hP, 2, 10, 15, 17
hP, 12
hP_expected (expected), 6
1rtest, 13
model.offset, 8, 11
na.exclude, 8, 11
na.fail, 8, 11
na.omit, 8, 11
nloptr, 8, 9, 11
offset, 8, 11
options, 8, 11
```

```
par, 15
phP (hP), 12
plot.glm_CMP, 18
plot.glm_CMP (plots), 14
plot.glm_hP, 18
plot.glm_hP (plots), 14
plots, 14, 18
predict.glm_CMP, 15
predict.glm_hP, 16
print.summary.glm_CMP
         (summary.glm_CMP), 19
print.summary.glm_hP(summary.glm_hP),
         19
residuals, 11, 17
rhP (hP), 12
summary, 9, 11
\verb|summary.glm_CMP|, 19|
\verb|summary.glm_hP|, 19|
```