Package 'pglm'

July 23, 2025

Version 0.2-3

Date 2021-07-19

Title Panel Generalized Linear Models

Depends R (>= 2.10), maxLik, plm

Imports statmod, Formula

Suggests lmtest, car

Description Estimation of panel models for glm-like models:
this includes binomial models (logit and probit), count models (poisson and negbin)
and ordered models (logit and probit), as described in:
Baltagi (2013) Econometric Analysis of Panel Data, ISBN-13:978-1-118-67232-7,
Hsiao (2014) Analysis of Panel Data <doi:10.1017/CBO9781139839327> and
Croissant and Millo (2018), Panel Data Econometrics with R, ISBN:978-1-118-94918-4.

License GPL (>= 2)

URL https://cran.r-project.org/package=pglm

NeedsCompilation no

Author Yves Croissant [aut, cre]

Maintainer Yves Croissant <yves.croissant@univ-reunion.fr>

Repository CRAN

Date/Publication 2021-07-19 18:10:02 UTC

Contents

`Fairness` *Perveived Fairness of Rules for Allocating Seats in Trains and Parking Spaces*

Description

observations of 401 individuals

number of observations : 5614

country : France

economic topic : public economics

econometrics topic : ordered response

Usage

```
data(Fairness)
```

Format

A dataframe containing :

id the individual index

answer a factor with levels 0 (very unfair), 1 (essentially unfair), 2 (essentially fair) and 3 (very fair)

good one of `'tgv'` (French fast train) and `'Parking'`

rule the allocation rule, a factor with levels `'peak'`, `'admin'`, `'lottery'`, `'addsupply'`, `'queuing'`, `'moral'` and `'compensation'`

driving does the individual has the driving license ?

education does the individual has a diploma ?

recurring does the allocation problem is reccuring ?

Source

provided by the authors.

References

Charles Raux, Stephanie Souche and Yves Croissant (2009) “How fair is pricing perceived to be? An empirical study”, *Public Choice*, **139(1)**, 227-240.

HealthIns *Health Insurance and Doctor Visits*

Description

observations of 401 individuals

number of observations : 20186

country : United States

economic topic : Health Economics

econometrics topic : censored dependant variable

Usage

```
data(HealthIns)
```

Format

A time serie containing :

id the individual index

year the year

med medical expenses

mdu number of face-to face medical visits

coins coinsurance rate

disease count of chronic diseases

sex a factor with level `'male'` and `'female'`

age the age

size the size of the family

child a factor with levels `'no'` and `'yes'`

Source

Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler and A. Leibowitz (1987) "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment", *American Economic Review*, **77(3)**, 251-277.

Deeb P. , and P.K. Trivedi (2002) "The structure of demand for medical care: latent class versus two-part models", *Journal of Health Economics*, **21**, 601-625..

References

http://cameron.econ.ucdavis.edu/musbook/mus.html.

PatentsRD *Patents, R\&d and Technological Spillovers for a Panel of Firms*

Description

annual observations of 181 firms from 1983 to 1991

number of observations : 1629

country : world

economic topic : producer behavior

econometrics topic : count data

Usage

```
data(PatentsRD)
```

Format

A dataframe containing :

firm firm's id

year year

sector firm's main industry sector, one of aero (aerospace), chem (chemistry), comput (computer), drugs, elec (electricity), food, fuel (fuel and mining), glass, instr (instruments), machin (machinery), metals, other, paper, soft (software), motor (motor vehicules)

geo geographic area, one of eu (European Union), japan, usa, rotw (rest of the world)

patent numbers of European patent applications

rdexp log of R and D expenditures

spil log of spillovers

Source

Cincer, Michele (1997) "Patents, R \& D and technological spillovers at the firm level : some evidence from econometric count models for panel data", *Journal of Applied Econometrics*, **12(3)**, may–june, 265–280.

References

Journal of Applied Econometrics data archive : http://qed.econ.queensu.ca/jae/.

PatentsRDUS *Dynamic Relation Between Patents and R\&d*

Description

yearly observations of 346 production units

number of observations : 3460

country : United States

economic topic : industrial economics

econometrics topic : count data

Usage

```
data(PatentsRDUS)
```

Format

A dataframe containing :

cusip compustat's identifying number for the firm

year year

ardssic a two-digit code for the applied R&D industrial classification

scisect is the firm in the scientific sector ?

capital72 book value of capital in 1972

sumpat the sum of patents applied for between 1972-1979

rd R&D spending during the year (in 1972 dollars)

patents the number of patents applied for during the year that were eventually granted

Source

Hall, Browyn, Zvi Griliches and Jerry Hausman (1986) "Patents and R and D: Is there a Lag?", *International Economic Review*, **27**, 265-283.

References

http://cameron.econ.ucdavis.edu/racd/racddata.html, chapter 9..

pglm *Panel Estimators for Generalized Linear Models*

Description

Estimation by maximum likelihood of glm (binomial and Poisson) and 'glm-like' models (Negbin and ordered) on longitudinal data

Usage

```
pglm(formula, data, subset, na.action,
     effect = c("individual", "time", "twoways"),
     model = c("random", "pooling", "within", "between"),
     family, other = NULL, index = NULL, start = NULL, R = 20,  ...)
```

Arguments

formula	a symbolic description of the model to be estimated,
data	the data: a pdata.frame object or an ordinary data.frame,
subset	an optional vector specifying a subset of observations,
na.action	a function which indicates what should happen when the data contains 'NA's,
effect	the effects introduced in the model, one of "individual", "time" or "twoways",
model	one of "pooling", "within", "between", "random",,
family	the distribution to be used,
other	for developper's use only,
index	the index,
start	a vector of starting values,
R	the number of function evaluation for the gaussian quadrature method used,
...	further arguments.

Value

An object of class "pglm", a list with elements:

coefficients	the named vector of coefficients,
logLik	the value of the log-likelihood,
hessian	the hessian of the log-likelihood at convergence,
gradient	the gradient of the log-likelihood at convergence,
call	the matched call,
est.stat	some information about the estimation (time used, optimisation method),
freq	the frequency of choice,
residuals	the residuals,

`fitted.values`	the fitted values,
`formula`	the formula (a `mFormula` object),
`expanded.formula`	the formula (a `formula` object),
`model`	the model frame used,
`index`	the index of the choice and of the alternatives.

Author(s)

Yves Croissant

Examples

```
## an ordered probit example
data('Fairness', package = 'pglm')
Parking <- subset(Fairness, good == 'parking')
op <- pglm(as.numeric(answer) ~ education + rule,
           Parking[1:105, ],
           family = ordinal('probit'), R = 5, print.level = 3,
           method = 'bfgs', index = 'id',  model = "random")

## a binomial (probit) example
data('UnionWage', package = 'pglm')
anb <- pglm(union ~ wage + exper + rural, UnionWage, family = binomial('probit'),
            model = "pooling",  method = "bfgs", print.level = 3, R = 5)

## a gaussian example on unbalanced panel data
data(Hedonic, package = "plm")
ra <- pglm(mv ~ crim + zn + indus + nox + age + rm, Hedonic, family = gaussian,
           model = "random", print.level = 3, method = "nr", index = "townid")

## some count data models
data("PatentsRDUS", package="pglm")
la <- pglm(patents ~ lag(log(rd), 0:5) + scisect + log(capital72) + factor(year), PatentsRDUS,
           family = negbin, model = "within", print.level = 3, method = "nr",
           index = c('cusip', 'year'))
la <- pglm(patents ~ lag(log(rd), 0:5) + scisect + log(capital72) + factor(year), PatentsRDUS,
           family = poisson, model = "pooling", index = c("cusip", "year"),
           print.level = 0, method="nr")

## a tobit example
data("HealthIns", package="pglm")
HealthIns$med2 <- HealthIns$med / 1000
HealthIns2 <- HealthIns[-2209, ]
set.seed(2)
subs <- sample(1:20186, 200, replace = FALSE)
HealthIns2 <- HealthIns2[subs, ]
la <- pglm(med ~ mdu + disease + age, HealthIns2,
           model = 'random', family = 'tobit', print.level = 0,
           method = 'nr', R = 5)
```

`UnionWage` *Unionism and Wage Rate Determination*

Description

yearly observations of 545 individuals from 1980 to 1987

number of observations : 4360

country : United States

economic topic : labor economics

econometrics topic : endogeneity

Usage

```
data(UnionWage)
```

Format

A dataframe containing :

id the individual index

year the year

exper the experience, computed as age - 6 - schooling

health does the individual has health disability ?

hours the number of hours worked

married is the individual married ?

rural does the individual lives in a rural area ?

school years of schooling

union does the wage is set by collective bargaining

wage hourly wage in US dollars

sector one of agricultural, mining, construction, trade, transportation, finance, businessrepair, personalservice, entertainment, manufacturing, pro.rel.service, pub.admin

occ one of proftech, manoffpro, sales, clerical, craftfor, operative, laborfarm, farmlabor, service

com one of black, hisp and other

region the region, one of NorthEast, NothernCentral, South and other

Source

Journal of Applied Econometrics data archive : http://qed.econ.queensu.ca/jae/.

References

Vella, F. and M. Verbeek (1998) "Whose wages do unions raise ? A dynamic model of unionism and wage", *Journal of Applied Econometrics*, **13**, 163–183.

Index