Package 'CalibInc'

September 27, 2017

Title Prediction of incidence using calibration model and proxy of incidence

Version 0.1

Author Edouard Chatignoux [aut,cre], Zoé Uhry [aut, ctb], Laurent Remontet [aut, ctb]

Maintainer Edouard Chatignoux <edouard.chatignoux@santepubliquefrance.fr>

Description
The package provides fonctions to predict incidence from proxy data using a calibration approach

Depends R (>= 3.3.3)

License GPL-3

Encoding UTF-8

LazyData true

Imports tidyverse,
scales,
splines,
RColorBrewer,
Matrix,
maptools

Suggests lme4,
MASS,
splines

RoxygenNote 6.0.1

R topics documented:

CalibInc   *Prediction of incidence from proxy data using a calibration approach*

Description

CalibInc functions takes a calibration model and a table with proxy data as arguments and returns a table with predictions and standard errors

Usage

```
CalibInc(mod, pred = NULL, weight = 1, aggregate = NULL,
  keep.Vp = FALSE)
```

Arguments

mod	Calibration model (type H/I ratio), fitted with glmer or glmmPQL
pred	Data frame containing proxy data to be calibrated. If NULL, try to use the data.frame used for model evaluation (searched in .GlobalEnv).
weight	Weight vector (size 1 for an uniform weight, same size as nrow(pred) otherwise). weight can be given as the name of a column in pred.
aggregate	Aggregate the predictions according to a combination of variables given as a formula (see examples). If NULL aggregation is done according to the combinations of all co-variables specified in the model formula.
keep.Vp	Set to TRUE to keep variances covariance matrix of the predictions

Value

A data frame containing predictions values (column pred) and se (column se) at the levels defined in aggregate. If keep.Vp=TRUE, the variances covariance matrix of the predictions is also returned as an attribute.

Author(s)

Edouard Chatignoux

Examples

```
library(lme4)
library(splines)
data(lopm.CalibSet)
data(lopm.Fr)
## Run the calibration model for lop cancer incidence
## and hospitalisation data in men, 2007-2011
k<-Hmisc::wtd.quantile(lopm.CalibSet$age,weights=lopm.CalibSet$C,p=0.5)%>%as.numeric
form.calib<-substitute(H~offset(log(C))+
                       ns(age,knots=k, Boundary.knots = range(age)+c(5,-5))+
                       (1|dist),
                       list(k=k))%>%as.formula
mod.calib<-glmer(form.calib,data=lopm.CalibSet%>%filter(C>0),family="poisson", nAGQ =20)
## Predict the total number of incident cases by district
CalibInc(mod.calib,pred=lopm.Fr,aggregate=~dist)
## Predict the total number of incident cases by age and district
CalibInc(mod.calib,pred=lopm.Fr,aggregate=~dist+age)
```

---

ggMap   *Plot variables values on a map*

Description

ggMap create a map from a data frame with observations on spatial units. It allows to represent values in continuous (with possible lower and upper bounds) or categorial scale, and to hatch non significant areas.

Usage

```
ggMap(var, data = NULL, map = NULL, limits = NULL, breaks = NULL,
  legend = list(breaks = NULL, density = F, title = NULL, pretty = T,
  placement = NULL), color = "Greys", rev = FALSE, na.action = na.pass,
  ns = NULL, path = TRUE)
```

Arguments

var	Variable name from the data `data` to map, or vector sorted in the order of the map spatial units
data	Data with column to plot. Must contain an `id` variable that identifies the district (jointure variable with `map` data)
map	Map shape ("polygg" type, set to ShapDep if NULL)
limits	Set limits if a continuous map is chosen
breaks	Values to cut var into classes
legend	List of legend parameters
	breaks Breaks in the legend
	density Set to `TRUE` to represent the legend as the density distribution of the observed values
	title Legend title (variable name otherwise)
	pretty If var is cut in classe with `cut` function, convert classes names to pretty labels
	placement Lengend placement (x0,x1,y0,y1) coordinates, in proportion of the plot zone
color	Name of the brewer palette [brewer.pal](#)
rev	Reverse color palette order (`TRUE`,`FALSE`)
na.action	What to do with missing values
ns	Variable name in the data set that identifies non significant areas (`TRUE` when non significant). In that case, the corresponding geographical areas are hatched.
path	Draw geographical boundaries (default to `TRUE`)

Value

A ggplot2 graph

Author(s)

Edouard Chatignoux

Examples

```
## Represent crude hospitalisation rate in France by district for LP cancer in men, 2007-2011
data(lopm.Fr)
crude.rate<-lopm.Fr%>%group_by(dist)%>%summarise(rate=100000*sum(H)/sum(py))%>%mutate(id=dist)
ggMap(var=rate,data=crude.rate,color= "YlOrRd",legend=list(title="Hosp. rate\nfor 100 000"))+
ggtitle("Crude hospitalisation rate for LOP, men, 2007-2011")
```

LogNormPI   *Confidence interval for log-normal distributed observations*

Description

LogNormPI calcultes confidence interval for log-normal distributed observations

Usage

```
LogNormPI(data, pred = pred, se = se, level = 0.95)
```

Arguments

data	A data frame with log-normal observations
pred	Log-normal observation
se	Standard error of pred
level	Confidence interval level

Author(s)

Edouard Chatignoux

Examples

```
dt<-data_frame(m=2,sd=1)
LogNormPI(data=dt,m,sd)
```

---

lopm.CalibSet   *Data on LOP cancer in men over the 2007-2011 period for the 14 districts covered by a cancer registry.*

Usage

```
data( lopm.CalibSet )
```

Format

A data frame with 182 rows and 5 variables

Details

- `dist` District
- `age` Central age of the age class
- `C` Number of cancer incident cases
- `H` Number of hospitalizations
- `py` Number of person-years

The number of cancer incident cases, corresponding numbers of hospitalizations and person-years are tabulated by 5 years age groups (from 0-5 to 90+, variable 'age' corresponds to the central age of the class) and district (variable 'dist'). In order to diminish the number of age classes with no cancer, the number of incident cases in lower ages classes were aggregated.

---

lopm.Fr	*Data on hospitalisation for LOP cancer in men over the 2007-2011 period for the 96 French adminsitrative districts*

Description

- `dist` District
- `age` Central age of the age class
- `H` Number of hospitalizations
- `py` Number of person-years
- `WSP` World Standard Population (Segi 1960 standard)

Usage

```
data( lopm.Fr )
```

Format

A data frame with 1728 rows and 5 variables

Details

The number of hospitalizations and person-years are tabulated by 5 years age groups (from 0-5 to 90+, variable 'age' corresponds to the central age of the class) and district (variable 'dist').

References

http://seer.cancer.gov/stdpopulations/

---

poly2ggplot   *Convert a spatialpolygon to a data frame for a ggplot2 use*

Description

Convert a spatialpolygon to a data frame for a ggplot2 use

Usage

```
poly2ggplot(file = NULL, poly = NULL, id = NULL)
```

Arguments

`file`	: File path to shape file
`poly`	: names of a SpatialPolygon if `NULL` file
`id`	: id variable for polygons

Value

A "polygg" object, to be used with ggplot2 (ggMap for example)

Author(s)

Edouard Chatignoux

ShapDep   *French administrative districts limits (Departement borders), simpli-*
                               fyied to be used with ggMap (polygg type)

Usage

```
data( ShapDep )
```

Format

A data frame with 2404 rows and 8 variables

Details

- `long` numeric (123255.6089–1242403.8507)
- `lat` numeric (6049568.1597–7110958.5478)
- `order` numeric (1–38)
- `hole` logical (0–0)
- `piece` factor avec 1 niveaux (1–1)
- `id` dept number : character (01–95)
- `group` factor avec 96 niveaux (01.1–95.1)
- `INSEE_DEP` character (01–95)

Index