Package 'wgeesel'

October 12, 2022

Type Package

Title Weighted Generalized Estimating Equations and Model Selection

Version 1.5

Date 2018-03-06

Encoding UTF-8

Author Cong Xu <congxu17@gmail.com>, Zheng Li <zheng.li@outlook.com>, Ming Wang <mwang@phs.psu.edu>

Maintainer Zheng Li <zheng.li@outlook.com>

Description Weighted generalized estimating equations (WGEE) is an extension of generalized linear models to longitudinal clustered data by incorporating the correlation within-cluster when data is missing at random (MAR). The parameters in mean, scale correlation structures are estimated based on quasi-likelihood. Multiple model selection criterion are provided for selection of mean model and working correlation structure based on WGEE/GEE.

Depends R (>= 3.0.0)

Imports MASS, stats, geepack, bindata, PoisNor,CRTgeeDR

NeedsCompilation no

Repository CRAN

Date/Publication 2018-03-09 05:08:21 UTC

R topics documented:

`wgeesel-package` *Weighted Generalized Estimating Equations and Model Selection*

Description

Weighted Generalized estimating equations (WGEE) is an extension of generalized linear models to longitudinal or clustered data by incorporating the correlation within-cluster when data is missing at random (MAR). The parameters in mean, scale, correlation structures are estimate based on quasi-likelihood. The package **wgeesel** also contains model selection criteria for variable selection in the mean model and for the selection of a working correlation structure in longitudinal data with dropout or monotone missingness using WGEE.

Details

The collection of functions includes:

`wgee` estimates parameters based on WGEE in mean, scale, and correlation structures, through mean link, scale link, and correlation link.

`QIC.gee`, `MQIC.gee`, `RJ.gee` calculate the QIC (QIC_u), MQIC (MQIC_u), Rotnitzky-Jewell criteria for variable selection in the mean model and/or selection of a working correlation structure in GEE (unbalanced data is allowed).

`MLIC.gee`, `QICW.gee` calculate the MLIC (MLICC) and QICW_r (QICW_p) for variable selection in the mean model and the selection of a working correlation structure in WGEE, which can accommodate dropout missing at random (MAR).

`data_sim` can simulate longitudinal response data in different distribution (gaussian, binomial, poisson) with drop missingness.

For a complete list of functions, use `library(help = "wgeesel")`.

Author(s)

Cong Xu <congxu17@gmail.com>, Zheng Li <zheng.li@outlook.com>, Ming Wang <mwang@phs.psu.edu>

Maintainer: Zheng Li <zheng.li@outlook.com>

References

Liang, K.Y. and Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, pp.13-22.

Preisser, J.S., Lohman, K.K. and Rathouz, P.J., 2002. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. *Statistics in medicine*, 21(20), pp.3035-3054.

Robins, J.M., Rotnitzky, A. and Zhao, L.P., 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90(429), pp.106-121.

Shen, C. W., & Chen, Y. H. (2012). Model selection for generalized estimating equations accommodating dropout missingness. *Biometrics*, 68(4), 1046-1054.

Wang, M., 2014. Generalized Estimating Equations in Longitudinal Data Analysis: A Review and Recent Developments. *Advances in Statistics*, 2014.

See Also

GEE methods exist for `geeglm` (**geepack**)

Examples

```
data(imps)

fit <- wgee(Y ~ Drug+Sex+Time,data=imps,id=imps$ID,family="binomial",
            corstr="exchangeable",scale=NULL,mismodel= R ~ Drug+Time)
```

`data_sim` *Simulate longitudinal data*

Description

The function generate correlated normal, Bernoulli or Poisson longitudinal data.

Usage

```
data_sim(id, rho, phi, x, beta, x_mis, para, corstr, family, lag_level)
```

Arguments

`id`	subject id
`rho`	with-in cluster correlation.
`phi`	scale parameter in the variance covariance matrix.
`x`	covariate associated with the response.
`beta`	coefficients associated with x.
`x_mis`	covariates associated with missing model.

para	coefficients associated with x_mis.
corstr	a character string specifies the working correlation structure. The following are permitted: "independence", "exchangeable","ar1".
family	a description of the error distribution and link function to be used in the model. This is a character string naming a family function. The following are permitted: "gaussian", "binary", "poisson".
lag_level	specify how many lags of response y will affect missing probability.

Details

Generate normal, poisson and binary longitudinal data based on the specified variance-covariance matrix.

Value

data	simulated dataset
prob_miss	missing percentage of the response y

Author(s)

Cong Xu, Zheng Li and Ming Wang

Examples

```
n=500
id=rep(1:n,each=3)
rho=1
phi=1
x=cbind(1,rnorm(length(id)))
beta=c(1,1)
x_mis=cbind(1,rnorm(length(id)))
para=c(1,1,1)
data_sim(id,rho,phi,x,beta,x_mis,para,"independent","gaussian",lag_level = 1)
```

---

drgee — *Fit Doubly Robust Weighted Generalized Estimating Equations*

Description

Analyzes longitudinal data with doubly robust augmented GEE approach.

Usage

```
drgee(model,outcomemodel, data, id, family, corstr, scale = NULL,
       mismodel = NULL, nameTRT, maxit=200, tol=0.001)
```

Arguments

`model`	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted.
`outcomemodel`	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the ouctome model
`data`	a data frame containing the variables in the model.
`id`	a vector which identifies the clusters. The length of "id" should be the same as the number of observations. Data are assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.
`family`	a description of the error distribution and link function to be used in the model. This is a character string naming a family function. The following are permitted: "gaussian", "binomial", "poisson".
`corstr`	a character string specifies the working correlation structure. The following are permitted: "independence", "exchangeable", "ar1","unstructured".
`nameTRT`	name of the variable containing information for the treatment
`scale`	a numeric variable giving the value to which the scale parameter should be fixed; if NA, the scale parameter is not fixed.
`mismodel`	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the missingness model to be fitted.
`maxit`	maximum iteration number for Newton-Raphson algorithm.
`tol`	the tolorance for the Newton-Raphson algorithm to converge.

Details

Analyzes longitudinal data with doubly robust augmented GEE approach.

Value

An object of type 'CRTgeeDR'

Author(s)

Zheng Li, Cong Xu and Ming Wang

References

Augmented GEE for improving efficiency and validity of estimation in cluster randomized trials by leveraging cluster-and individual-level covariates -2012-Stephens A., Tchetgen Tchetgen E. and De Gruttola V. : *Stat Med* 31(10)-915-930.

See Also

https://cran.r-project.org/web/packages/CRTgeeDR/index.html

Examples

```
data(imps)
fit <- drgee(IMPS79 ~ Drug+Sex+Time, IMPS79 ~ Drug+Sex+Time,
      data=imps, id=imps$ID, family="gaussian",nameTRT="Drug",
       corstr="exchangeable", scale=NULL, mismodel= R ~ Drug+Time)
```

---

imps *Inpatient Multidimensional Psychiatric Scale (IMPS)*

Description

The imps data frame has 1544 rows and 8 columns. The data is from National Institute of the Mental Health Schizophrenia Collaborative Study, where the effect of chlorpromazine, fluphenazine, or thioridazine treatment on the overall severity of the schizophrenia disorder is of interest.

Usage

```
data(imps)
```

Format

This data frame contains the following columns:

ID patient ID

IMPS79 the severity of the schizophrenia disorder (ranges from 0 to 7)

Week the fixed visit time

Drug the indicator of treatment (1: chlorphromazine, fluphenazine, or thioridazine treatment; 0: placebo)

Sex the indicator of sex of the patients (1: male; 0: female)

R an indicator of the missingness (1: observed; 0: missing)

Time square root of the Week covariate

Y an indicator if IMPS $>= 4$

Source

Gibbons, R.D. and Hedeker, D., 1994. Application of random-effects probit regression models. *Journal of consulting and clinical psychology*, 62(2), p.285.

References

Shen, C. W., & Chen, Y. H. (2012). Model selection for generalized estimating equations accommodating dropout missingness. *Biometrics*, 68(4), 1046-1054.

MLIC.gee *MLIC and MLICC for Weighted GEE*

Description

Calculate the MLIC (missing longitudinal information criterion) for selection of mean model, and the MLICC (missing longitudinal information correlation criterion) for selection of working correlation structure, based on the expected quadratic loss and the WGEE.

Usage

```
MLIC.gee(object,object_full)
```

Arguments

`object`	a fitted model object of class `"wgee"`.
`object_full`	a fitted model object of class `"wgee"`: the largest candidate model under consideration to be fitted.

Value

Return a data frame of MLIC, MLICC and Wquad_loss.

Note

MLIC and MLICC model selection criterion for longitudinal data criterion with dropouts or monotone missingness under the assumption of MAR.

Author(s)

Cong Xu, Zheng Li and Ming Wang

References

Robins, J.M., Rotnitzky, A. and Zhao, L.P., 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90(429), pp.106-121.

Shen, C.W. and Chen, Y.H., 2012. Model selection for generalized estimating equations accommodating dropout missingness. *Biometrics*, 68(4), pp.1046-1054.

Shen, C.W. and Chen, Y.H., 2013. Model selection of generalized estimating equations with multiply imputed longitudinal data. *Biometrical Journal*, 55(6), pp.899-911.

See Also

`wgee`

Examples

```
data(imps)
fit1 <- wgee (Y ~ Drug+Sex+Time,data=imps,id=imps$ID,family="binomial",
             corstr="exchangeable",scale=NULL,mismodel= R ~ Drug+Time)

fit_f <- wgee (Y ~ Drug+Sex+Time+Time*Sex+Time*Drug,data=imps,id=imps$ID, family="binomial",
              corstr="exchangeable",scale=NULL,mismodel= R ~ Drug+Time)
###not run#####
##MLIC.gee(fit1,fit_f)
```

MQIC.gee *MQIC and MQICu for GEE*

Description

Calculate MQIC (Modified QIC) and MQIC$_u$. MQIC is an asymptotic unbiased estimator of the risk function based on the independent quasi-likelihood. MQIC and original QIC may have non-negligible effect for model selection, especially when the true correlation structure completely different from the working correlation structure.

Usage

```
MQIC.gee(object)
```

Arguments

`object`	a fitted model object of class `"wgee"`.

Value

Return a list of MQIC, MQIC$_u$ and Quasi-likelihood.

Author(s)

Cong Xu, Zheng Li and Ming Wang

References

Pan, W., 2001. Akaike's information criterion in generalized estimating equations. *Biometrics*, 57(1), pp.120-125.

Gosho, M., Hamada, C., and Yoshimura, I. (2011). Modifications of QIC and CIC for Selecting a Working Correlation Structure in the Generalized Estimating Equation Method. *Japanese Journal of Biometrics*, 32(1), 1-12.

Imori, S., 2013. On Properties of QIC in Generalized Estimating Equations. *Hiroshima University*, pp.1-8.

See Also

geeglm (**geepack**)

Examples

```
data(imps)

fit <- wgee(IMPS79 ~ Drug+Sex+Time, data=imps,id=imps$ID,family="gaussian",
             corstr="independence")
MQIC.gee(fit)


data(ohio)

fit2 <- wgee(resp ~ age + smoke+age:smoke, data=ohio, id=ohio$id,
             family="binomial", corstr="independence")
MQIC.gee(fit2)
```

---

ohio — *Ohio Children Wheeze Status*

Description

The ohio data frame has 2148 rows and 4 columns. The dataset is a subset of the six-city study, a longitudinal study of the health effects of air pollution.

Usage

```
data(ohio)
```

Format

This data frame contains the following columns:

resp an indicator of wheeze status (1=yes, 0=no)

id a numeric vector for subject id

age a numeric vector of age, 0 to 9 years old

smoke an indicator of maternal smoking at the first year of the study (1=yes, 0=no)

References

Fitzmaurice, G.M. and Laird, N.M. (1993) A likelihood-based method for analyzing longitudinal binary responses, *Biometrika* **80**: 141–151.

See Also

ohio (**geepack**)

QIC.gee *QIC and QICu for GEE*

Description

Calculate quasi-likelihood under the independence model criterion (QIC) and QIC_u based on GEE.

Usage

```
QIC.gee(object)
```

Arguments

`object`	a fitted model object of class "wgee".

Value

Return a vector of QIC, QIC_u and Quasi-likelihood.

Note

QIC can be used to select the best correlation structure and the best fitting model in GEE analyses. The GEE is fitted by `geeglm` (**geepack**). QIC_u is a simplified version of QIC, which can not be applied to select the optimal working correlation structure. `geeglm` (**geepack**) only works for complete data. Thus if there are NA's in data, the missing values are automatically removed by `na.omit`.

Author(s)

Cong Xu, Zheng Li and Ming Wang

References

Liang, K.Y. and Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, pp.13-22.

Pan, W., 2001. Akaike's information criterion in generalized estimating equations. *Biometrics*, 57(1), pp.120-125.

Prentice, R.L. and Zhao, L.P., 1991. Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics*, pp.825-839.

See Also

`geeglm` (**geepack**). **MuMIn** also provides QIC value.

Examples

```
data(imps)
fit <- wgee(Y ~ Drug+Sex+Time, data=imps, id=imps$ID, family="binomial",
            corstr="exchangeable")

QIC.gee(fit)

data(seizure)
###reshapre the seizure data to "long" format
seiz.long <- reshape(seizure, varying=list(c("base","y1", "y2", "y3", "y4")),
                      v.names="y", times=0:4, direction="long")
seiz.long <- seiz.long[order(seiz.long$id, seiz.long$time),]

fit <- wgee(y ~ age + trt + time, data=seiz.long, id=seiz.long$id,
            family="poisson", corstr="independence")

QIC.gee(fit)
```

QICW.gee *QICWr and QICWp for WGEE*

Description

Calculate the QICW_r and QICW_p (an information criterion based on the weighted quasi-likelihood function) for selection of mean model and correlation structure based on the WGEE.

Usage

```
QICW.gee(object)
```

Arguments

`object`	a fitted model object of class "wgee".

Value

Return a data frame of QICW_r, QICW_p and Wquasi_lik.

Note

QICW_r can be used for variable selection and for selecting the correlation structure in WGEE analyses. QICW_p is a simplified version of QICW_r, which can not be applied to select the optimal working correlation structure in WGEE.

Author(s)

Cong Xu, Zheng Li and Ming Wang

References

Gosho, M., 2015. Model selection in the weighted generalized estimating equations for longitudinal data with dropout. *Biometrical Journal*.

Platt, R.W., Brookhart, M.A., Cole, S.R., Westreich, D. and Schisterman, E.F., 2013. An information criterion for marginal structural models. *Statistics in medicine*, 32(8), pp.1383-1393.

Robins, J.M., Rotnitzky, A. and Zhao, L.P., 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90(429), pp.106-121.

See Also

wgee

Examples

```
data(imps)

### variable selection by QICWr, not rum###
#fit <- wgee(IMPS79 ~ Drug+Sex+Time, mismodel= R ~ Drug+Time, data=imps,
##           id=imps$ID, family="gaussian", corstr="exchangeable")
##QICW.gee(fit)

#fit <- wgee(IMPS79 ~ Drug+Sex+Time+Time:Sex+Time:Drug, mismodel= R ~ Drug+Time,
#       data=imps, id=imps$ID, family="gaussian", corstr="exchangeable")
##QICW.gee(fit)
```

RJ.gee *RJC for GEE*

Description

Calculate RJC (Rotnitzky-Jewell information criterion) based on GEE.

Usage

```
RJ.gee(object)
```

Arguments

object	a fitted model object of class "wgee".

Details

Rotnitzky-Jewell information criterion (RJC) is usually used for working correlation structure selection.

Value

Return the value of the Rotnitzky-Jewell information criterion (RJ).

Author(s)

Cong Xu, Zheng Li and Ming Wang

References

Rotnitzky, A. and Jewell, N.P., 1990. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika*, pp.485-497.

See Also

`geeglm` (**geepack**)

Examples

```
data(ohio)

fit <- wgee(resp ~ age + smoke+age:smoke, data=ohio, id=ohio$id,
            family="binomial", corstr="exchangeable")

RJ.gee(fit)
```

RJ2.gee *Corrected RJC for GEE*

Description

Calculate corrected RJC (Rotnitzky-Jewell information criterion) based on GEE with a modified robust variance estimator.

Usage

```
RJ2.gee(object)
```

Arguments

`object`	a fitted model object of class `"wgee"`.

Value

Return the value of the corrected Rotnitzky-Jewell information criterion (RJC).

Note

RJ2.gee can only handle balanced data (data with dropout missingness). Two assumptions should be satisfied. (A1) The conditional variance of Y_{ij} given X_{ij} is correctly specified; (A2) A common correlation structure, Rc, exists across all subjects. If there is missingness, one can group the subjects by the cluter size of the response variable. And, calculate the modified robust variance in each group to get the pooled estimate of the variance.

Author(s)

Cong Xu, Zheng Li and Ming Wang

References

Rotnitzky, A. and Jewell, N.P., 1990. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika*, pp.485-497.

Wang, M. and Long, Q., 2011. Modified robust variance estimator for generalized estimating equations with improved small-sample performance. *Statistics in Medicine*, 30(11), pp. 1278-1291.

See Also

geeglm (**geepack**), RJ.gee

Examples

```
data(ohio)

fit <- wgee(resp ~ age + smoke+age:smoke, data=ohio, id=ohio$id,
            family="binomial",corstr="exchangeable")
RJ2.gee(fit)
```

---

seizure *Epiliptic Seizures*

Description

The dataset has the number of epiliptic seizures in each of four two-week intervals,and in a baseline eight-week inverval, for treatment and control groups with a total of 59 individuals.

Usage

```
data(seizure)
```

Format

This data frame contains the following columns:

y1 the number of epiliptic seizures in the 1st 2-week interval

y2 the number of epiliptic seizures in the 2nd 2-week interval

y3 the number of epiliptic seizures in the 3rd 2-week interval

y4 the number of epiliptic seizures in the 4th 2-week interval

trt an indicator of treatment

base the number of epilitic seizures in a baseline 8-week interval

age a numeric vector of subject age

Source

Thall, P.F. and Vail S.C. (1990) Some covariance models for longitudinal count data with overdispersion. *Biometrics* **46**: 657–671.

References

Diggle, P.J., Liang, K.Y., and Zeger, S.L. (1994) Analysis of Longitudinal Data. Clarendon Press.

See Also

`seizure` (**geepack**)

`wgee` *Fit Weighted Generalized Estimating Equations (WGEE)*

Description

`wgee` fits weighted generalized estimating equations (WGEE) with Newton Raphson algorithm. `wgee` has a syntax similar to `glm` and returns an object similar to a `glm` object.

Usage

```
wgee(model, data, id, family, corstr, scale = NULL, mismodel = NULL, maxit=200, tol=0.001)
```

Arguments

`model`	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted.
`data`	a data frame containing the variables in the model.
`id`	a vector which identifies the clusters. The length of "id" should be the same as the number of observations. Data are assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula.

family	a description of the error distribution and link function to be used in the model. This is a character string naming a family function. The following are permitted: "gaussian", "binomial", "poisson".
corstr	a character string specifies the working correlation structure. The following are permitted: "independence", "exchangeable", "ar1","unstructured".
scale	a numeric variable giving the value to which the scale parameter should be fixed; if NA, the scale parameter is not fixed.
mismodel	an object of class "formula" (or one that can be coerced to that class): a symbolic description of the missingness model to be fitted.
maxit	maximum iteration number for Newton-Raphson algorithm.
tol	the tolorance for the Newton-Raphson algorithm to converge.

Details

wgee analyzes longitudinal data with missing values by weighted genralized estimating equations (WGEE), proposed by Robins, Totnizky and Zhao (1995). WGEE can handle missing at random problem. The standard error of the estimates are calculated as described in (Fitzmaurice, Laird, and Ware, 2011) and Preisser, Lohman, and Rathouz (2002).

Value

beta	covariate effect estimates
var	variance covariances estimates for beta
w_r_square	weighted R square for continuous data
mu_fit	fitted values of response
scale	scale estimates
rho	rho estimates
weight	The weight of response y
model	WGEE model structure
x	covariates in WGEE
y	response in WGEE
mis_fit	esimates of the missingness model
call	the function to be called
id	as input
data	as input
family	as input
corstr	as input

Author(s)

Zheng Li, Cong Xu and Ming Wang

References

Fitzmaurice, G.M., Laird, N.M. and Ware, J.H., 2012. *Applied longitudinal analysis (Vol. 998)*. John Wiley & Sons.

Liang, K.Y. and Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. *Biometrika*, pp.13-22.

Preisser, J.S., Lohman, K.K. and Rathouz, P.J., 2002. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. *Statistics in medicine*, 21(20), pp.3035-3054.

Robins, J.M., Rotnitzky, A. and Zhao, L.P., 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90(429), pp.106-121.

Rubin, D.B., 1976. Inference and missing data. *Biometrika*, pp.581-592.

See Also

`geeglm` (**geepack**)

Examples

```
####Example1
data(imps)

fit <- wgee(IMPS79 ~ Drug+Sex+Time, data=imps, id=imps$ID, family="gaussian",
            corstr="exchangeable", scale=NULL, mismodel= R ~ Drug+Time)

#####Example2
data(seizure)

###reshapre the seizure data to "long" format
seiz.long <- reshape(seizure,
                     varying=list(c("base","y1", "y2", "y3", "y4")),
                     v.names="y", times=0:4, direction="long")
seiz.long <- seiz.long[order(seiz.long$id, seiz.long$time),]

###create missing value for seiz.long dataset
set.seed(12345)
obs <- exp(9+seiz.long$age*(-0.2))/(1+exp(9+seiz.long$age*(-0.2)))
R <- lapply(unique(seiz.long$id),function(x){
  idx=which(seiz.long$id==x)
  r=c()
  r[1]=1
  for(j in 2:length(idx)){
    if(r[j-1]==1){
      r[j]=rbinom(1,1,obs[idx[j]])
    }
    else r[j]=0
  }
  return(r)
})
remove_id <- which(sapply(R,sum)==1)
```

```
remove_idx <- which(seiz.long$id %in% remove_id==1)
seiz.long <- cbind(seiz.long,R=unlist(R))[-remove_idx,]
seiz.long$y_mis <- ifelse(seiz.long$R,seiz.long$y,NA)

###fit WGEE, not run ###
fit <- wgee(y_mis ~ age + trt + time, data=seiz.long, id=seiz.long$id,family="poisson",
            corstr="exchangeable",scale=NULL, mismodel= R ~ age)
```

`WRsquare.gee` *Weighted R Square for WGEE*

Description

Calculate the weighted R^2 (missing longitudinal information criterion) for selection of mean model.

Usage

```
WRsquare.gee(object, weight_mean)
```

Arguments

`object`	a fitted model object of class "wgee".
`weight_mean`	logical; whether weighted mean of the response should be used for weighted R square.

Value

Return a list of weighted R square.

Author(s)

Zheng Li, Cong Xu and Ming Wang

References

Nelder, J.A. and Baker, R.J., 1972. *Generalized linear models.* Encyclopedia of statistical sciences.

See Also

wgee

Examples

```
data(imps)

fit <- wgee(Y ~ Drug+Sex+Time, data=imps, id=imps$ID, family="binomial",
            corstr="exchangeable",scale=NULL,mismodel= R ~ Drug+Time)
WRsquare.gee(fit, weight_mean=TRUE)
```

ylag *Generate subject-level lagged response*

Description

Suppose you have longitudinal response y and it's subject id. This function generates lagged y for each subject.

Usage

```
ylag(id,y,lag,na=FALSE)
```

Arguments

id	subject id
y	response
lag	how many lags for y
na	logical; whether remain NAs in the response when lag > 1. Default to FALSE and output 0s instead of NAs.

Value

return a vector of lagged y

Author(s)

Cong Xu, Zheng Li and Ming Wang

Examples

```
id <- rep(c(1:20),each=3)
y <- rnorm(length(id))
ylag(id,y,1) #lag=1
ylag(id,y,2,na=FALSE) #lag=2
```

Index