Package 'mreg'

October 26, 2023

Type Package

Title Fits Regression Models When the Outcome is Partially Missing

Version 1.2.1

Date 2023-10-26

Description Implements the methods described in Bond S, Farewell V, 2006, Exact Likelihood Estimation for a Negative Binomial Regression Model with Missing Outcomes, Biometrics.

License GPL (>= 3)

Encoding UTF-8

Language en-GB

LazyData true

URL https://github.com/shug0131/mreg

RoxygenNote 7.2.3

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Simon Bond [aut, cre] (<https://orcid.org/0000-0003-2528-1040>), Vern Farewell [aut]

Maintainer Simon Bond <simon.bond7@nhs.net>

Repository CRAN

Date/Publication 2023-10-26 16:30:04 UTC

R topics documented:

mreg *To perform regression when discrete outcome variables are missing*

Description

This software was created for the paper referred to below. If a longitudinal data base has regularly updated explanatory variables, but whose outcome variable is only intermittently collected then we can still perform exact maximum likelihood estimation of a regression model if the outcome variable is discrete.

Usage

```
mreg(
  formula,
  data,
  patid,
  start.theta = NULL,
  modify = unity,
  modify.p = 0,
  mod.formula = ~1,
  density.name = "negbin",
  link = "log",
  iterlim = 100,
  gradtol = 1e-06,
  steptol = 1e-06,
  na.action = NULL,
  print.level = 2,
  zero.start = FALSE
)
```

Arguments

formula	This is a formula object e.g. Y~A+B to describe the location parameter
data	This is a data frame in which the variables are recorded
patid	In a longitudinal context this indexes the individuals. Note that the observations within each patient is assumed to be ordered according the timing of the observations.
start.theta	Optional vector of starting values for location and nuisance parameters
modify	We may wish to let the location depend on functions of the previous outcomes. Since these may be missing, we have to provide a function that can cope with all the potential values the outcome may have taken. See paper
modify.p	This is the dimension of the parameters associated with the modify function.
mod.formula	If we require other variables to interact with the previous observation we must create a set of variables to use. This is a one-sided formula e.g. ~X+Z, if we wanted to use those variables.

density.name	This is the density the increment in outcome is assumed to follow. It can be one of three values: negbin, poisson, geometric.
link	This is the link function $g(\mu) = \eta$. Where η is a linear combination of covariates, and μ is the expected value of the outcome. The link function can be one of four values: identity, log, logit, hyper.
iterlim	The maximum number of iterations allowed for the nlm function.
gradtol	The parameter gradtol for the nlm function.
steptol	The parameter steptol for the nlm function
na.action	Parameter is not used: If any covariates are missing the function will return an error.
print.level	The parameter print.level for the nlm function. Set to the maximum, verbose level.
zero.start	It may be the case that it is known that the first value of the outcome was zero for all individuals, in which case invoke this TRUE/FALSE option.

Value

It returns an object of class mreg which is similar to a lm object. It has print and summary methods to display the fitted parameters and standard errors.

References

Bond S, Farewell V, 2006, Exact Likelihood Estimation for a Negative Binomial Regression Model with Missing Outcomes, Biometrics

See Also

print.mreg, summary.mreg, paper, unity

Examples

```
data(public)
## Not run:
mod1 <- mreg( damaged~offset(log(intervisit.time))+esr.init,
data=public,patid=ptno,print.level=2, iterlim=1000 )
mod.ncar <-mreg(damaged ~ offset(log(intervisit.time)) + esr.init +
         tender + effused + clinic.time, data = public, patid = ptno,
         modify = paper, modify.p = 5, mod.formula = ~art.dur.init,
       density.name = "negbin.ncar", iterlim = 1000, print.level = 2)

## End(Not run)
```

`mreg_package` *mreg: Implements the techniques of exact likelihood when the discrete outcome can be missing in a regression mode*

Description

Implements the methods described in Bond S, Farewell V, 2006, Exact Likelihood Estimation for a Negative Binomial Regression Model with Missing Outcomes, Biometrics, Submitted. The main function is `mreg`.

References

Bond S, Farewell V, 2006, Exact Likelihood Estimation for a Negative Binomial Regression Model with Missing Outcomes, Biometrics

---

`paper` *An example of a function to calculate the effect on the location that depends on previous outcome in an mreg model*

Description

In a regression model one may want to let the linear predictor depend on previous values of the outcome variable in longitudinal data. When the outcome variable is missing we can still do this but we have to create a function that calculates a vector of linear predictors with one element for each of the possible preceding values of the outcome.

Usage

```
paper(x, y, mod.Z)
```

Arguments

`x`	is a vector of possible values the previous value of the response could take.
`y`	is an vector of the coefficients. Its length is specified in the `mreg` function by the argument `modify.p`.
`mod.Z`	vector of observed covariates that may interact with the unobserved preceding outcome corresponding to the observation. It is taken from the `model.frame` produced by the `mod.formula` argument in `mreg`.

Details

This is an example function that was used in the paper referred to below. The user may write their own function as long as it takes the arguments specified above and returns a vector the same length as `x` that will be added to the linear predictor. It is up to the user to ensure that their function identifies the correct column of `mod.Z` using indices (`[]` that correspond to the desired variables in `mod.formula`. If any function other than `unity` (which does nothing to the linear predictor) is used then it is not possible to produce fitted values or residuals.

Value

A numeric vector the same length as `x` that will be added to the linear predictor. It also has two subsidiary attributes: `par.names` names to be used to label the associated coefficients, and `par.dim` the length of this vector of coefficients.

See Also

`unity` `mreg`

Examples

```
## The function is currently defined as
function(x,y, mod.Z){
  #x is the imputed response
  #y is the set of parameters
 #mod.Z is a VECTOR/matrix of explanatory variables
 rad.type <- cut(x, breaks=c(-1,0,4,9,50))
 if( is.vector(mod.Z)){
   arthdur.first <- rep(mod.Z[2],length(x))
 }
 else{
   arthdur.first <- rep(mod.Z[1,2], length(x))
 }
 X <- model.matrix( ~rad.type+I(x==0):arthdur.first)
 structure( X[,-1, drop=FALSE]%*%y, par.names= colnames( X)[-1],par.dim=dim(X)[2]-1)
}
```

`print.mreg` *Prints the coefficients from an mreg object*

Description

Prints the coefficients from an mreg object

Usage

```
## S3 method for class 'mreg'
print(x, digits = max(3, getOption("digits") - 3), ...)
```

Arguments

`x`	an mreg object
`digits`	number of digits to print decimals to.
`...`	other arguments, not currently used but required for generic methods

Value

It prints the coefficients from an `mreg` regression model.

See Also

mreg summary.mreg

---

public *Subset of Arthritis Data*

Description

This is a subset of 100 patients and variables from the clinical data used in the paper below. *Not to be used in publications without permission of the authors.*

@format A data frame with 943 observations on the following 8 variables.

`damaged` The number of radiologically damaged joints

`intervisit.time` Time between clinic visits

`esr.init` The first ESR measurement

`art.dur.init` The duration of arthritis at first clinic visit

`tender` Count of the number of tender joints

`effused` Count of the number of effused joints

`ptno` Anonymous patient number

`clinic.time` Total time spend under clinical care

Usage

```
public
```

Format

An object of class `data.frame` with 943 rows and 8 columns.

Source

Bond S, Farewell V, 2006, Exact Likelihood Estimation for a Negative Binomial Regression Model with Missing Outcomes

summary.mreg *Prints a summary of an mreg object*

Description

@description Similar to summary.glm, it prints the coefficients, standard errors, Wald tests, residuals (if available) and dispersion from an mreg object.

Usage

```
## S3 method for class 'mreg'
summary(
  object,
  digits = max(3, getOption("digits") - 3),
  symbolic.cor = object$symbolic.cor,
  signif.stars = getOption("show.signif.stars"),
  ...
)
```

Arguments

object	An mreg object
digits	the number of significant digits to use when printing.
symbolic.cor	logical. If TRUE, print the correlations in a symbolic form (see symnum) rather than as numbers.
signif.stars	logical. If TRUE, 'significance stars' are printed for each coefficient.
...	further arguments passed to or from other methods.

Value

prints to output a description of the fitted model.

See Also

mreg print.mreg

unity *The default value of 'modify' function in 'mreg'*

Description

If the location term in a regression model does not depend on any previously observed values of the outcome in a longitudinal data set, then we obtain simplification in our estimation procedure when the outcomes can be missing. Using the default value of `unity` for the argument `modify` in the `mreg` function does this.

Usage

```
unity(x, y, mod.Z)
```

Arguments

`x`	is a vector of possible values the previous value of the response could take.
`y`	is an vector of the coefficients. Its length is specified in the `mreg` function by the argument `modify.p`.
`mod.Z`	vector of observed covariates that may interact with the unobserved preceding outcome corresponding to the observation. It is taken from the `model.frame` produced by the `mod.formula` argument in `mreg`.

Details

This function is the default value for the argument `modify` for `mreg`. It does nothing to the linear predictor term. For this function `unity` there are no such covariates. A default value for `mod.formula` is `~1`.

Value

A vector of zeroes the same length is the argument `x`.

Index