Package 'isni'

February 29, 2020

Title Index of Local Sensitivity to Nonignorability

Version 1.2

Author Hui Xie <huixie@uic.edu>, Weihua Gao, Baodong Xing, Daniel Heitjan, Donald Hedeker, Chengbo Yuan

Description The current version provides functions to compute, print and summarize the Index of Sensitivity to Nonignorability (ISNI) in the generalized linear model for independent data, and in the marginal multivariate Gaussian model and the mixed-effects models for continuous and binary longitudinal/clustered data. It allows for arbitrary patterns of missingness in the regression outcomes caused by dropout and/or intermittent missingness. One can compute the sensitivity index without estimating any nonignorable models or positing specific magnitude of nonignorability. Thus ISNI provides a simple quantitative assessment of how robust the standard estimates assuming missing at random is with respect to the assumption of ignorability. For a tutorial, download at <http://huixie.people.uic.edu/Research/ISNI_R_tutorial.pdf>. For more details, see Troxel Ma and Heitjan (2004) and Xie and Heitjan (2004) <doi:10.1191/1740774504cn005oa> and Ma Troxel and Heitjan (2005) <doi:10.1002/sim.2107> and Xie (2008) <doi:10.1002/sim.3117> and Xie (2012) <doi:10.1016/j.csda.2010.11.0

Depends R (>= 3.0.1)

Imports nlme, mvtnorm, nnet, matrixcalc, Formula, mixor, lme4

License GPL-2

Encoding UTF-8

LazyData true

Maintainer Hui Xie <huixie@uic.edu>

Suggests knitr, rmarkdown

RoxygenNote 7.0.2

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2020-02-29 07:30:02 UTC

R topics documented:

---

coc   *A data set for Psychiatric Drug Treatment*

Description

The variables are as follows:

- y. number of days
- time. in weeks
- sub. patients id
- group. treatment group with placebo (PBO) and desipramine (DMI)
- basey. baseline days
- g1. dummy variables for treatment group
- g2. dummy variables for treatment group
- time1. dummy variables for time
- time2. dummy variables for time
- grptime. group by time

Usage

```
data(coc)
```

Format

A data frame with 869 rows and 10 variables

---

definemissingstatus	*Utility function to generate missing status variables in longitudinal data with dropout and/or intermittent missingness.*

Description

Utility function to generate missing status variables in longitudinal data with dropout and/or intermittent missingness.

Usage

```
definemissingstatus(data, id, time, y)
```

Arguments

data	the name of the panel dataset in the long format with each row denoting a subject-visit observation for ALL the planned visits, regardless of being missed or not. When a subject is lost to follow up, the data set must include the observation at the first time of being lost to follow up.
id	the name of the level-2 clustering variable.
time	the name of the variable denoting the time of the visit. Can set time=NULL if data is already sorted by id and time within id.
y	the name of the outcome variable of the interest that is subject to missingness.

Value

a dataset with the following three new variables added:

- g_ : missingness indicator, "O"-observed, "I"-intermittent missing, "D"-dropout

- gp_: missingness indicator in the previous visit, "O"-observed, "I"-intermittent missing, "D"-dropout, "U"-undefined.

- yp_: the immediately observed prior outcome.

Examples

```
qolefnew <- definemissingstatus(qolef, id=id, time=time, y=y)
```

---

isniglm   *Function for ISNI computation when the outcome follows GLMs.*

Description

Calculate the ISNI when the regression outcome is subject to missingness and follows generalized linear models (GLMs)

Usage

```
isniglm(
  formula,
  family = gaussian,
  data,
  weights,
  subset,
  start = NULL,
  offset
)
```

Arguments

formula	an object of class "Formula": a symbolic description of the models to be fitted for the outcome and missingness status variable. The details of model specification are given under "Details".
family	a description of the error distribution to be used in the GLM for the outcome.
data	the name of data frame containing the variables in the model and all the observations including those intended to be collected but became missing.
weights	an optional vector of "prior weights" to be used in the fitting process for the outcome model and the missingness mechanism model. Should be NULL or a numeric vector.
subset	an optional vector specifying a subset of observations to be used in the fitting process for the outcome model and the missingness mechanism model.
start	starting values for the parameters in the linear predictor of the outcome model.
offset	an optional vector to specify an a priori known component to be included in the linear predictor during fitting the GLM for the outcome. This should be NULL or a numeric vector of length equal to the number of observations.

Details

The ISNI analysis is based on a joint selection model and requires specifying two model equations: the complete-data model and the missing data mechanism model. To specify the variables in the models and required for computing the ISNI measures, we make use of the R package "Formula" designed for handling model equations with multiple responses and multiple sets of predictors. At a minimum, the user should supply a single-equation in the typical form: response ~ Xterms

where `response` is the (numeric or factor) vector for the outcome of interest and `Xterms` is a series of terms, separated by + operators, which specify a linear predictor for `response`. With the signle-equation specification, the `isnimgm` function will by default use `(is.na(response))` as the missingness status variable and `Xterms` as the observed predictors for missingness. The `isniglm` then computes the MAR estimates and conducts ISNI computation to evaluate the rate of change of model estimates in the neighborhood of the MAR model where the missingness probability is allowed to depend on the unobserved value of `response`, even after conditioning on the other missingness predictors.

The above single-equation formula specification uses the same set of predictors for the outcome model and the missingness mechanism model for computing ISNI. To use different sets of predictors, one can explicitly specifiy a two-equation formula as: `response | is.na(response) ~ Xterms | Sterms`, which specifies the formula for the complete-data model as `response ~ Xterms` and that for the missing data mechanism model as `is.na(response) ~ Sterms`, and `Xterms` and the observed predictors for missingness `Sterms` can be different.

References

Troxel AB, Ma G, Heitjan DF. An Index of Local Sensitivity to Nonignorability. Stat Sin. 2004;14:1221-1237.

Examples

```
## load data set
data(sos)

## Perform the MAR analysis
ymodel= sexact  ~ gender*faculty
summary(glm(ymodel,family=binomial, data=sos))

## Perform ISNI analysis
sos.isni<-isniglm(ymodel, family=binomial, data=sos)
sos.isni
summary(sos.isni)

## specifying the missing data model explicitly
ygmodel= sexact | is.na(sexact)  ~ gender*faculty | gender *faculty
summary(isniglm(ygmodel, family=binomial, data=sos))


 ## ISNI for grouped binomial regression.
gender <- c(0,0,1,1,0,0,1,1)
faculty    <- c(0,0,0,0,1,1,1,1)
gender = factor(gender, levels = c(0, 1), labels =c("male", "female"))
faculty = factor(faculty, levels = c(0, 1), labels =c("other", "mdv"))

SAcount <- c(NA, 1277, NA, 1247, NA, 126, NA, 152)
total       <- c(1189,1710,978,1657,68,215,73,246)
sosgrp <- data.frame(gender=gender, faculty=faculty, SAcount=SAcount, total=total)
ymodel <- SAcount/total ~gender*faculty
sosgrp.isni<-isniglm(ymodel, family=binomial, data=sosgrp, weight=total)
```

---

isniglmmbin	*Function for ISNI computation when the longitudinal/clustered binary outcome follows a GLMM.*

Description

Calculate the ISNI when the regression outcome is subject to missingness and follows generalized linear mixed-effects models (GLMMs) for binary outcomes.

Usage

```
isniglmmbin(
  formula,
  data,
  random,
  id,
  weights,
  subset,
  predprobobs,
  misni = FALSE,
  method = 1
)
```

Arguments

formula	an object of class "Formula": a symbolic description of the models to be fitted for the outcome and missingness status variable. The details of model specification are given under "Details".
data	the name of data frame containing all the variables in the model and all the observations including those intended to be collected but became missing.
random	an object of class "formula": an one-sided linear formula description of the random-effects part of the model to be fitted for the outcome
id	the name of the level-2 clustering variable.
weights	frequency weights to be assigned to each id. when supplied, indicates differential weights are used; otherwise each id is weighted equally.
subset	an optional vector specifying a subset of observations to be used in the fitting process for the outcome model and the missingness mechanism model.
predprobobs	Null if using buil-in multinomial transitional logistic model to obtain predicted probabilities of being observed; otherwise user supply the name of the variable in data that gives these probabilities for all the observations in the data.
misni	FALSE if using the default approach to computing ISNI with a scalar nonignorability parameter; TRUE when computing ISNI with multiple nonignorability parameters.
method	Indicate the method to obtain the MAR estimates: 1: OPTIM(); 2: MIXOR(); 3: GLMER()

Details

The ISNI analysis is based on a joint selection model and requires specifying two model equations: the outcome model and the missingness mechanism model. At a minimum, the user should supply a single-equation in the typical form: `response ~ Xterms` where `response` is the (numeric or factor) vector for the binnary outcome of interest and `Xterms` is a series of terms, separated by + operators, which specify a linear predictor for response. With the signle-equation specification, the `isniglmm` function will by default use the utility function `definemissingstatus` provided in the package to generate the missingness status variables at the current and prior visits and then use `Xterms` as the predictors for fitting a first-order transitional missing data model. It is important to sort within-id observations by time so that the missingness status variables can be defined correctly in this default setting. The `isniglmmbin()` then computes the MAR estimates and conducts ISNI computation to evaluate the rate of change of model estimates in the neighborhood of the MAR model where the missingness probability is allowed to depend on the unobserved value of `response`, even after conditioning on the other missingness predictors.

The above single-equation formula specification uses the same set of predictors for the outcome model and the missingness mechanism model for computing ISNI. To use different sets of predictors, one can explicitly specifiy a two-equation formula as: `response | miss + missprior ~ Xterms | Sterms`, which specifies the formula for the outcome model as `response ~ Xterms` and that for the missingness mechanism model as `miss | missprior ~ Sterms`, where `Xterms` and `Sterms` can be different, `miss` and `missprior` are the variable names in `data` denoting the missingness status at the current and prior visits, respectively.

For `isniglmm`, `response ~ Xterms` specfied the fixed-effect part of the linear mixed-effects model for the outcome. The random-effect part of the model is specified as a one-sided formula via the argument `random`.

References

Xie H. A Local Sensitivity Analysis Approach to Longitudinal Non-Gaussian Data with Non-Ignorable Dropout. Statist Med. 2008;27:3155-3177.
Xie H. Analyzing Longitudinal Clinical Trial Data with Nonignorable Missingness and Unknown Missingness Reasons. Comput Stat Data Anal. 2012;56:1287-1300.
Xie H, Qian Y. Measuring the Impact of Nonignorability in Panel Data with Non-Monotone Non-response. Journal of Applied Econometrics. 2012;27:129-159.

Examples

```
data(skquit)

formula=quit~time
## formula=quit~time + helmert1:as.factor(time) +helmert2:as.factor(time)+ helmert3:as.factor(time)
random=~1

result=isniglmmbin(formula, skquit, random, id,misni=FALSE,method=2)
summary(result)
```

---

isnilmm   *Function for ISNI computation when the outcome follows LMM.*

Description

Calculate the ISNI when the regression outcome is subject to missingness and follows linear mixed-effects models (LMMs)

Usage

```
isnilmm(formula, data, random, id, weights, subset, predprobobs, misni = FALSE)
```

Arguments

formula   an object of class "Formula": a symbolic description of the models to be fitted
                for the outcome and missingness status variable. The details of model specifica-
                tion are given under "Details".

data   the name of data frame containing all the variables in the model and all the
                observations including those intended to be collected but became missing.

random   an object of class "formula": an one-sided linear formula description of the
                random-effects part of the model to be fitted for the outcome

id   the name of the level-2 clustering variable.

weights   frequency weights to be assigned to each id. when supplied, indicates differen-
                tial weights are used; otherwise each id is weighted equally.

subset   an optional vector specifying a subset of observations to be used in the fitting
                process for the outcome model and the missingness mechanism model.

predprobobs   Null if using buil-in multinomial transitional logistic model to obtain predicted
                probabilities of being observed; otherwise user supply the name of the variable
                in data that gives these probabilities for all the observations in the data.

misni   FALSE if using the default approach to computing ISNI with a scalar nonignor-
                ability parameter; TRUE when computing ISNI with multiple nonignorability
                parameters.

Details

The ISNI analysis is based on a joint selection model and requires specifying two model equa-
tions: the complete-data model and the missing data mechanism model. To specify the variables
in the models that are required for computing the ISNI measures, we make use of the R package
"Formula" designed for handling model equations with multiple responses and multiple sets of pre-
dictors . At a minimum, the user should supply a single-equation in the typical form: response ~
Xterms where response is the (numeric or factor) vector for the outcome of interest and Xterms is
a series of terms, separated by + operators, which specify a linear predictor for response. With
the signle-equation specification, the isniglm function will by default use the utility function
definemissingstatus provided in the package to generate the missingness status variables at the
current and prior visits and then use Xterms as the observed missingness predictors for fitting a

first-order transitional missing data model. It is important to sort within-`id` observations by time so that the missingness status variables can be defined correctly in this default setting. The `isnimgm` then computes the MAR estimates and conducts ISNI computation to evaluate the rate of change of model estimates in the neighborhood of the MAR model where the missingness probability is allowed to depend on the unobserved value of `response`, even after conditioning on the other missingness predictors.

The above single-equation formula specification uses the same set of predictors for the outcome model and the missingness mechanism model for computing ISNI. To use different sets of predictors, one can explicitly specifiy a two-equation formula as: `response | miss + missprior ~ Xterms | Sterms`, which specifies the formula for the complete-data model as `response ~ Xterms` and that for the missing data mechanism model as `miss + missprior ~ Sterms`, where `Xterms` and the observed predictors for missingness `Sterms` can be different, `miss` and `missprior` are the variable names in `data` denoting the missingness status at the current and prior visits, respectively.

For `isnilmm`, `response ~ Xterms` specfied the fixed-effect part of the linear mixed-effects model for the outcome. The random-effect part of the model is specified as a one-sided formula via the argument `random`.

References

Hui Xie and Yi Qian (2012) Measuring the impact of nonignorability in panel data with non-monotone nonresponse., Journal of Applied Econometrics 27: 129-159.
Hui Xie, Gao,W, Xing, B., Heitjan, D, Hedeker, D and Yuan, C. (2018) Measuring the Impact of Nonignorable Missingness Using the R packaeg isni, Computer Methods and Programs in Biomedicine 164 207-220.

Examples

```
data(qolef)
ymodel= y | g+ gp~   as.factor(time)*group+perf+sever

##Random intercept model
result=isnilmm(ymodel, random=~1, id=id,  data=qolef)
summary(result)
```

---

isnimgm	*Function for ISNI computation when the outcome follows marginal multivariate Gaussian Models.*

Description

Calculate the ISNI when the regression outcome is subject to missingness and follows marginal multivaraite Gaussian models.

Usage

```
isnimgm(
  formula,
  data,
  cortype = "CS",
  id,
  subset,
  weights,
  predprobobs,
  misni = FALSE
)
```

Arguments

formula	an object of class "Formula": a symbolic description of the models to be fitted for the outcome and missingness status variable. The details of model specification are given under "Details".
data	the name of data frame containing all the variables in the model and all the observations including those intended to be collected but became missing.
cortype	the type of within-subject correlation structure.
id	the name of variable for the level-2 clustering variable.
subset	an optional vector specifying a subset of observations to be used in the fitting process for the outcome model and the missingness mechanism model.
weights	frequency weights to be assigned to each id. when supplied, indicates differential weights are used; otherwise each id is weighted equally.
predprobobs	Null if using buil-in multinomial transitional logistic model to obtain predicted probabilities of being observed; otherwise user supply the name of the variable in data that gives these probabilities for all the observations in the data.
misni	FALSE if using the default approach to computing ISNI with a scalar nonignorability parameter; TRUE when computing ISNI with multiple nonignorability parameters.

Details

The ISNI analysis is based on a joint selection model and requires specifying two model equations: the complete-data model and the missing data mechanism model. To specify the variables in the models that are required for computing the isni measures, we make use of the R package "Formula" designed for handling model equations with multiple responses and multiple sets of predictors . At a minimum, the user should supply a single-equation in the typical form: response ~ Xterms where response is the (numeric or factor) vector for the outcome of interest and Xterms is a series of terms, separated by + operators, which specify a linear predictor for response. With the signle-equation specification, the isniglm function will by default use the utility function definemissingstatus provided in the package to generate the missingness status variables at the current and prior visits and then use Xterms as the observed missingness predictors for fitting a first-order transitional missing data model. It is important to sort within-id observations by time so that the missingness status variables can be defined correctly in this default setting. The isnimgm

then computes the MAR estimates and conducts ISNI computation to evaluate the rate of change of model estimates in the neighborhood of the MAR model where the missingness probability is allowed to depend on the unobserved value of response, even after conditioning on the other missingness predictors.

The above single-equation formula specification uses the same set of predictors for the outcome model and the missing data mechanism model for computing ISNI. To use different sets of predictors, one can explicitly specifiy a two-equation formula as: response|miss + missprior ~ Xterms | Sterms, which specifies the formula for the outcome model as response ~ Xterms and that for the missing data mechanism model as miss + missprior ~ Sterms, where Xterms and the observed missingness predictors Sterms can be different, miss and missprior are the variable names in data denoting the missingness status at the current and prior visits, respectively.

References

Ma G, Troxel AB, Heitjan DF. An Index of Local Sensitivity to Nonignorable Dropout in Longitudinal Modeling. Stat Med. 2005;24:2129-2150.
Xie H. Analyzing Longitudinal Clinical Trial Data with Nonignorable Missingness and Unknown Missingness Reasons. Comput Stat Data Anal. 2012;56:1287-1300.
Xie H, Qian Y. Measuring the Impact of Nonignorability in Panel Data with Non-Monotone Nonresponse. Journal of Applied Econometrics. 2012;27:129-159.

Examples

```
models= y | g+gp ~   perf + sever+ as.factor(time) + group +as.factor(time):group   |
      as.factor(time) * group + yp+ perf + sever
qolef.isni=isnimgm(models, data=qolef, id=id)
summary(qolef.isni)
```

---

print.isniglm   *Function to print the isniglm object.*

Description

print method for class isniglm

Usage

```
## S3 method for class 'isniglm'
print(x, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

x	the isniglm object obtained from the isniglm function
digits	the number of significant digits to use when printing
...	further arguments passed to or from other methods.

Value

   The function prints the model call, isni and c statistics from the isniglm object.

   `print.isniglmm`   *Function to print the isniglmm object.*

Description

   `print` method for class isniglmm

Usage

```
## S3 method for class 'isniglmm'
print(x, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

   x   the isniglmm object obtained from the isniglmm function

   digits   the number of significant digits to use when printing

   ...   further arguments passed to or from other methods.

Value

   The function print.isnilmm prints the model call, isni and c statistics from the isniglmm object.

   `print.isnilmm`   *Function to print the isnilmm object.*

Description

   `print` method for class isnilmm

Usage

```
## S3 method for class 'isnilmm'
print(x, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

   x   the isnilmm object obtained from the isnilmm function

   digits   the number of significant digits to use when printing

   ...   further arguments passed to or from other methods.

Value

   The function print.isnilmm prints the model call, isni and c statistics from the isnilmm object.

---

print.isnimgm   *Function to print the isnimgm object.*

Description

`print` method for class isnimgm

Usage

```
## S3 method for class 'isnimgm'
print(x, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

x	the isnimgm object obtained from the isnimgm function
digits	the number of significant digits to use when printing
...	further arguments passed to or from other methods.

Value

The function print.isnimgm prints the model call, isni and c statistics from the isnimgm object.

---

qolef   *A data set for Quality of Life Emontional Functioning outcome.*

Description

The variables are as follows:

- id: patients id
- y: EF score
- time: time in months since randomization.
- group: placebo (0) or flutamide (1)
- perf: baseline performance score
- sever: baseline disease severity
- basey: EF at baseline
- g: missingness status ("O"=observed, "D"=dropout, "I"=intermittent missingness)
- gp: missingness status in the prior visit ("O"=observed, "D"=dropout, "I"=intermittent missingness, "U"=undefined)

Usage

```
data(qolef)
```

Format

A data frame with 2860 rows and 10 variables

---

skquit *A randomzied trial data set for Smoking cessation*

Description

The variables are as follows:

- id. subjects id
- time. Time in weeks after baseline
- quit. 0: smoking; 1: quit smoking
- sub. patients id
- helmert1. Herlmert contrast 1 among 4 treatment groups.
- helmert2. Herlmert contrast 2 among 4 treatment groups.
- helmert3. Herlmert contrast 3 among 4 treatment groups.

Usage

```
data(skquit)
```

Format

A data frame with 1861 rows and 6 variables

---

sos *Dataset for a survey of sexual behavior*

Description

A dataset from a survey of sexual behavior

Usage

```
data(sos)
```

Format

A data frame with 6136 rows and 3 variables

Details

This data frame contains the following three factor variables:

- sexact: response to the survey question "Have you ever had sexual intercouse" with two levels of no (ref level) and yes.

- gender: two levels of male (ref level) and female.

- faculty: the student's faculty with two levels of mdv=medical/dental/veterinary (ref level) and other =all the other faculty categories.

---

summary.isniglm   *Function to print out a summary of isniglm object in a matrix form.*

Description

Function to print out a summary of isniglm object in a matrix form.

Usage

```
## S3 method for class 'isniglm'
summary(object, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

object	the isniglm object obtained from the isniglm function
digits	the number of significant digits to use when printing
...	further arguments passed to or from other methods.

Value

The function summarizes the MAR coefficient estimates, standard errors, isni and c statistics from the isniglm object as a matrix form.

---

summary.isniglmm   *Function to print out a summary of isniglmm object in a matrix form.*

Description

Function to print out a summary of isniglmm object in a matrix form.

Usage

```
## S3 method for class 'isniglmm'
summary(object, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

`object`	the isniglmm object obtained from the isniglmm function
`digits`	the number of significant digits to use when printing
`...`	additional arguments

`summary.isnilmm`	*Function to print out a summary of isnilmm object in a matrix form.*

Description

Function to print out a summary of isnilmm object in a matrix form.

Usage

```
## S3 method for class 'isnilmm'
summary(object, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

`object`	the isnilmm object obtained from the isnilmm function
`digits`	the number of significant digits to use when printing
`...`	additional arguements

`summary.isnimgm`	*Function to print out a summary of isnimgm object in a matrix form.*

Description

Function to print out a summary of isnimgm object in a matrix form.

Usage

```
## S3 method for class 'isnimgm'
summary(object, digits = max(3, getOption("digits") - 2), ...)
```

Arguments

`object`	the isnimgm object obtained from the isnimgm function
`digits`	the number of significant digits to use when printing
`...`	additional arguements

---

tmdm

Function to fit the transitional missing data model and obtain the predicted probabilities of being observed for all observations.

Description

Fit missing data model and obtain predicted probabilities of being observed for all observations.

Usage

```
tmdm(formula, data, weights, subset)
```

Arguments

formula	a formula to specify a multinomial transitional missing data model in the form of g+gp ~Sterms.
data	the name of the dataset for fitting missing data mechansim model
weights	an optional vector of "prior weights" to be used in the fitting process for the missingness mechanism model. Should be NULL or a numeric vector.
subset	an optional vector specifying a subset of observations to be used in the fitting process for the outcome model and the missingness mechanism model.

Examples

```
qolefitg <- tmdm(g+gp~as.factor(time)+group+perf+sever,data=qolef)
```

Index