Package 'mimix'

March 12, 2021

Title Reference-based imputation for longitudinal clinical trials with protocol deviation

Version 0.0.12

Author Kevin McGrath <kkevinmcgrath@yahoo.co.uk>

Maintainer Ian White Ian.White <Ian.White@ucl.ac.uk>

Description This package imputes missing numerical outcomes for a longitudinal trial with protocol deviations.
It uses distinct treatment arm-
based assumptions for the unobserved data, following the general algorithm
of Carpenter, Roger, and Ken-
ward (2013), and the causal model model of White, Royes and Best (2019).
Sensitivity analysis to departures from these assump-
tions can be done by the Delta method of Roger.
The program is derived from the mimix Stata package written by Suzie Cro, with additional coding for
the causal model and delta method.
The reference-based methods are jump to reference (J2R), copy increments in reference (CIR), copy reference (CR),
and the causal model, all of which must specify the reference treatment arm. Other methods are missing at random (MAR) and the last mean carried forward (LMCF). Individual-specific imputation methods
(and their reference groups) can be specified.

URL https://https://github.com/UCL/mimix

License GPL-3

Encoding UTF-8

LazyData true

Imports data.table, Hmisc, norm2, mice, pastecs

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

R topics documented:

`acupuncture` *Sample data: acupuncture trial*

Description

A data set containing results of a randomised, double-blind, parallel-group comparing active treatment with placebo The primary outcome is head, measured at time 3 and 12

Usage

```
acupuncture
```

Format

A data frame with 802 rows and 11 columns

id

time

age

sex

migraine

chronicity

practice_id

treat

head_base covariate

head dependent variable

withdrawal_reason

Examples

```
## Not run:
 impCausalref1 <- mimix(data=acupuncture,covar=c("head_base","sex"),depvar=head,treatvar=treat,
       idvar=id,timevar=time,
       method="Causal",reference=1,K0=1,K1=0.5,M=5,seed=54321)
 library(mice)
 fit<-with(as.mids(impCausalref1), lm(head.12~treat+head_base+sex))
 summary(pool(fit))

## End(Not run)
```

`AddDelta` *ddd delta's to imputed values*

Description

add delta's to imputed values

Usage

```
AddDelta(vec_tst, ncovar, mata_imp, delta, dlag)
```

Arguments

`vec_tst`	vector of visit names
`ncovar`	number of covariates
`mata_imp`	the imputed values (as well as the complete) and missing pattern
`delta`	vector (a values in Roger's paper) length = number of time points
`dlag`	vector (b values in Roger's paper) length = number of time points

Details

Adding delta values after wthdrawal Specifying delta and dlag allows imputations to differ sytematically from RBI methods. They provide an increment which is added on to all values imputed after treatment discontinuation, but not to interim (intermediate) missing values. Values of delta are cumulated after treatment discontinuation. For example, for an individual who discontinued treatment at the 2nd time point, we take the vector of delta's starting at the 3rd time point and add their cumulative sums to the imputed values. Specifying dlag modifies this behaviour, so that the vector of delta's starting at the 3rd time point is multipled elementwise by the vector dlag. The formula for the increment at time k for an individual who discontinued after time p is b_1xa_p+1 + b_2xa_p+2 + ... + b_k-pxa_k where delta=(a_1,a_2,...) and dlag=(b_1,b_2,...). A common increment of 3 at all time points after treatment discontinuation is achieved by setting delta=c(3,3,3,...) and dlag=c(1,0,0,...).

Value

mata_imp the adjusted imputed vaues (and unadjusted non-missing)

analyselist *find descriptive stats on the M imputed data set*

Description

find descriptive stats on the M imputed data set

Usage

```
analyselist(id, datlist, varlist)
```

Arguments

id	patient identifier
datlist	imputed dataset of M imputations
varlist	list of derived variables ,varlist <- c("fev.2","fev.4","fev.8","fev.12","base")

Details

select on patient id and find their means etc

Value

printout of descriptve stats

Examples

```
## Not run:
varlist <- c("fev.2","fev.4","fev.8","fev.12","base")
analyselist(5099,impdataset,varlist)

## End(Not run)
```

antidepressant *Sample data: antidepressant trial*

Description

A data set containing antidepressant trial data as described in paper by White,Royes,Best (2019) The primary outcome is HAMD17.TOTAL measured at visit number 4,5,6,7.

Usage

```
antidepressant
```

Format

dataframe containing 688 rows and 14 columns

PATIENT.NUMBER
HAMA.TOTAL
PGI_IMPROVEMENT
VISIT...VISIT.3.DATE
VISIT.NUMBER
TREATMENT.NAME
PATIENT.SEX
POOLED.INVESTIGATOR
basval
HAMD17.TOTAL dependent variable
change
miss_flag
methodcol individual-specific method
referencecol individual-specific reference arm

Examples

```
## Not run:
 impIndiv <- mimix(data=antidepressant,covar=c("basval","PATIENT.SEX"),depvar=HAMD17.TOTAL,
 treatvar=TREATMENT.NAME,idvar=PATIENT.NUMBER,
 timevar=VISIT.NUMBER,methodvar="methodcol",referencevar="referencecol",M=5,seed=54321)
 library(mice)
 fit<-with(data= as.mids(impantiIndivDt),lm(HAMD17.TOTAL.7~TREATMENT.NAME+basval+PATIENT.SEX))
 summary(pool(fit))

## End(Not run)
```

---

asthma *Sample data: asthma trial*

Description

A data set containing asthma trial data as used in the Stata mimix help file The primary outcome variable is fev, measured at 2,4,8,12 weeks

Usage

```
asthma
```

Format

A data frame containing 732 rows and 5 columns

id patient identifier
time
treat
base covariate
fev dependent variable

Examples

```
## Not run:
 impJ2Rridge<-(mimix(data=asthma,covar=c("base"),depvar=fev,treatvar=treat,idvar=id,timevar=time,
    method="J2R",reference=1,delta=c(0.5,0.5,1,1 ),M=5,seed=101,prior=ridge,burnin=1000)
 library(mice)
 fit<-with(data= as.mids(impJ2Rridge),lm(fev.12~treat))
 summary(pool(fit))

## End(Not run)
```

`Causal_loop` *process Causal method*

Description

process Causal method

Usage

```
Causal_loop(c_mata_miss, mata_Means, MeansC, K0, K1)
```

Arguments

`c_mata_miss`	vector of col locaton of missing values , eg 5 6
`mata_Means`	vector of means after mcmc draws eg 17 1 16.8 15.5 14.6 13.2
`MeansC`	vector of means after mcmc draws using variance from reference group
`K0`	Causal constant for use with Causal method
`K1`	exponential decaying Causal constant for use with Causal method 0<k1<1

Details

This is based on "White,Royes,Best" paper

Value

mata_means

`CIR_loop` *process CIR method*

Description

process CIR method

Usage

```
CIR_loop(c_mata_miss, mata_Means, MeansC)
```

Arguments

c_mata_miss	vector of col locaton of missing values , eg 5 6
mata_Means	vector of means after mcmc draws eg 17 1 16.8 15.5 14.6 13.2
MeansC	vector of means after mcmc draws using variance from reference group

Details

This is based on Suzie Cro's Stata program

Value

mata_means

---

fillinterims *fills missing interims distinguishing from post-discontinuation*

Description

fills missing interims (ie prior to post-discontinuation) assuming MAR

Usage

```
fillinterims(impdata, interims, Mimp = M, idvar)
```

Arguments

impdata	the imputed data under MAR for m=0 to M
interims	the interim id's
Mimp	the number of imputations specified , ie M total imputsations
idvar	the patient id

Details

checks methodindiv not null

Value

list of 1st data set with interims imputed plus M interim cases of each interim case to be matche in 2nd pass

getimpdatasets *to obtain the M'th imputed data set from the output list into one dataset*

Description

to append the M imputed data sets wit hte original unimputed data

Usage

```
getimpdatasets(varlist)
```

Arguments

varlist	list of data containing imputed values from the M pattern groups

Details

This combines the imputations found from the M pattern groups

Value

impdatasets

ifmethodindiv *performs imputation for individual-specific method*

Description

alternative logic for individual method

Usage

```
ifmethodindiv(
  methodvar,
  referencevar,
  mg,
  m,
  M,
  paramBiglist,
  i,
  treatvar,
  c_mata_nonmiss,
  c_mata_miss,
  mata_miss,
  mata_nonmiss,
  K0,
  K1
)
```

Arguments

methodvar	individual method col
referencevar	individual reference col
mg	pattern lookup table
m	where we are in the imputations
M	number of total imputations.
paramBiglist	list of Beta and Sigma parameters from mcmc
i	in loop through mg rows
treatvar	treatment group
c_mata_nonmiss	vector of positions of nonmissing
c_mata_miss	2,3,4 vector of missing positionals
mata_miss	0,1 indicators of missing values in repeated time visits
mata_nonmiss	0,1 indicators of nonmissing values
K0	Causal constant for use with Causal method
K1	exponential decaying Causal constant for use with Causal method

Details

checks methodindiv not null

Value

list of outputs

LMCF_loop *process LMCF method*

Description

process LMCF method

Usage

```
LMCF_loop(c_mata_miss, mata_Means)
```

Arguments

c_mata_miss	vector of col locaton of missing values , eg 5 6
mata_Means	vector of means after mcmc draws eg 17 1 16.8 15.5 14.6 13.2

Details

This is based on Suzie Cro's Stata program

when no observed data first mean is used in Stata

may be different here

Value

mata_means

`mimix` *Main function for performing reference-based multiple imputation of longitudinal data*

Description

main wrapper for running mimix

similar to the Stata mimix function

Usage

```
mimix(
  data,
  covar,
  depvar,
  treatvar,
  idvar,
  timevar,
  method = NULL,
  reference = NULL,
  methodvar = NULL,
  referencevar = NULL,
  K0 = 1,
  K1 = 1,
  delta = NULL,
  dlag = NULL,
  M = 1,
  seed = 101,
  prior = jeffreys,
  burnin = 1000,
  bbetween = NULL,
  mle = FALSE
)
```

Arguments

`data`	Dataset in long format
`covar`	Covariates - baseline. Must be complete (no missing values), enclose in quotes.
`depvar`	Dependent (outcome) variable
`treatvar`	Treatment group, can be character
`idvar`	Participant identifier.
`timevar`	Time point for repeated measure
`method`	Reference-based imputation method: must be J2R, CR, CIR, MAR, Causal or LMCF, enclose in quotes.
`reference`	Reference group for J2R, CIR, CR methods
`methodvar`	column in data-set specifying individual method, enclose in quotes.
`referencevar`	column in data-set specifying reference group as forindividual method, enclose in quotes.

K0	Causal constant for use with Causal method
K1	exponential decaying Causal constant for use with Causal method
delta	vector of delta values to add onto imputed values (non-mandatory) (a's in Five_Macros user guide),length as number of time points
dlag	vector of delta values to add onto imputed values (non-mandatory) (b's in Five_Macros use guide),length as number of time points
M	Number of imputations to be created
seed	Seed value. Specify this so that a new run of the command will give the same imputed values.
prior	Prior when fitting multivariate normal distributions> It can be one of jeffreys (default), uniform or ridge
burnin	Number of burn-in iterations when fitting multivariate normal distributions.
bbetween	Number of iterations between imputed data sets when fitting multivariate normal distributions.
mle	logical optionlibrary(mice) to Use maximum likelihood parameter estimates instead of MCMC draw parameters

Details

The program works through the following steps

1. set up a summary table based on treatment arm and missing data pattern

(i.e. which timepoints are unobserved)

1. Fit a multivariate normal distribution to each treatment sarm using MCMC methods in package norm2

1. Impute all interim missing values under a MAR assumption, looping ove treatments and patterns

1. Impute post-discontinuation missing values under the user-specified assumption,

looping over treatments and patterns (and over methodvar and referncevar if specified)

1. Perform delta-adjustment if specified

1. Repeat steps 2-5 M times and form into a single data frame

The baseline value of the outcome could be handed as an outcome, but this would allow a treatment effect at baseline

We instead recommend handling it as a covariate

The program is based on Suzie Cro's Stata program mimix

The user can use the as.mids() function in the mice package to convert the output data to mids data type and hence

to perform analysis using Rubin's rules.

The mimix package contains the functions preprodata and preproIndivdata to process long longitudinal data into wide data format

pass2Loop performs 2nd pass after interims found by MAR

Also the function Addelta to add delta adjustment to the imputed estimates

Value

The M imputed data sets are output concatenated as one large data frame appended to the original unimputed data

References

Roger J. (2017) MissingmacroDoc38_DIAversion.doc

Examples

```
## Not run:
"performing jump to reference with treatment reference arm 1 on asthma trial data"
mimixout<-mimix(data=asthma,covar=c("base"),depvar=fev,treatvar=treat,idvar=id,timevar=time,
 method="J2R", reference=1,M=5,seed=54321)
library(mice)
"Fitting regression model to find treatment effects using Rubin's rules by
     treating output data frame as.mids() object "
fit<-with(data= as.mids(mimixout),expr = lm(fev.12~treat+base))
summary(pool(fit))
mimix(data=acupuncture,covar= c("head_base"),depvar=head,treatvar=treat,idvar=id,
 timevar=time,method="CIR",reference=1,M=5,seed=54321,
  prior=jeffreys,burnin=1000)

## End(Not run)
```

---

mimix_Comparison *mimix: Comparisons with Stata and SAS*

Description

mimix is available in Stata whils th Five_macros suite does similar in SAS

Comparison with Stata

This mimix is based on the Stata version and has similar functionality while adding the causal method and delta adjustment. As with the Stata version the input data requires the longitudinal input data in long format with one record per individual at each timepoint. The program differs in how interim missing cases - those cases which have a missing measurement at a timepoint previous to a later observed measurement - are treated. Under Stata by default, the interim missing are treated the same as for the post-discontinuation missing unless the interim option is explicitly used. Here the interims are treated as under MAR, the post-discontinuations then imputed under the specified method. There is no interim option as there is in Stata. Unlike Stata an option is supplied whereby the prior used in the MCMC draws can be changed from the default jeffreys (as in Stata) to either the ridge or uniform

Comparison with SAS

Whilst this program is based on the Stata program, the latter is an adaptation of the SAS macro miwithd, written by james Roger, subsequently updated to the Five_Macros suite of macros This program uses the same approach for the delta adjustment as described in the Five_macros, in comparing outputs from our program with the Five_macros it is to be noted that interaction between

treatment and covariates is not allowed in the SAS macros, and comparisons are only valid for example in testing the Causal model by specifically not not using the covbytime and catcovbytime options in the Five_macros Not using these options also means that the LMCF method can be compared with either ALMCF or OLMCF in the Five_macros. When there is no observed data (common in the acupuncture data) the first mean is used in Stata, a warning is given in the Five_macros

References

Cro s, Morris T, Kenward G,Carpenter Joshttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC5796638/
White I, Royes J, Best N, https://arxiv.org/abs/1705.04506
Roger J,
URL: https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#sensitivity-analysis,

---

`pass2Loop` *Performs the imputation for the specified method after MAR ran*

Description

2nd pass for specified method after 1st pass MAR ran

Usage

```
pass2Loop(
  Imp_Interims,
  method,
  mg,
  ntreat,
  depvar,
  covar,
  treatvar,
  reference,
  trtgp,
  mata_Obs,
  mata_all_newlist,
  paramBiglist,
  idvar,
  flag_indiv,
  M,
  delta,
  dlag,
  K0,
  K1
)
```

Arguments

`Imp_Interims`	Interim cases
`method`	• Specified model to run Reference-based imputation method
`mg`	the summary table based on missing data patern
`ntreat`	vector of treatment groups

depvar	response variable
covar	covariate variable(s)
treatvar	Treatment group, coded 1,2,..
reference	Reference group for J2R, CIR, CR methods
trtgp	treatmet grp
mata_Obs	raw data with interims imputed
mata_all_newlist	raw data with interims imputed in list
paramBiglist	list of MCMC beta and Sigma parameters
idvar	Participant id
flag_indiv	flag whether specified individual column in data
M	number of imputations
delta	vector of delta values to add onto imputed values (non-mandatory) (a's in Rogers paper),length as number of time points
dlag	vector of delta values to add onto imputed values (non-mandatory) (b's in Rogers paper),length as number of time points
K0	Causal constant for use with Causal method
K1	exponential decaying Causal constant for use with Causal method

Details

reads the summary table based on missing data pattern- mg mimix_group

reflects the pattern and treatment group configuration of the raw data

then acts as a looping mechanism, norm2 is used as MCMC multivariate normal

Value

impdataset the M imputed data-sets appended to the "missing values" data-set in wide format

Examples

```
## Not run:
testpass2impdatset<- pass2Loop(Imp_Interims,method,mg,ntreat,depvar,treatvar,reference,trtgp,mata_Obs,mata_
paramBiglist,idvar,flag_indiv,M,delta,K0,K1)

## End(Not run)
```

---

preprodata *pre-process long longitudinal data into wide format*

Description

process data into wide format for group method

Usage

```
preprodata(
  data,
  covar,
  depvar,
  treatvar,
  idvar,
  timevar,
  M,
  reference,
  method = NULL
)
```

Arguments

data	data in long format
covar	covariates and base depvar complete
depvar	dependent variable
treatvar	treatment group
idvar	patient id
timevar	time variable for repeated visit
M	number imputations
reference	reference group
method	RBI method

Details

checks method finds missingness pattern

Value

list of outputs

---

preproIndivdata *pre-process long longitudinal data into wide format for individual-specific*

Description

process data into wide format for individual-specified method

Usage

```
preproIndivdata(
  data,
  covar,
  depvar,
  treatvar,
  idvar,
```

```
  timevar,
  M,
  reference = NULL,
  method = NULL,
  methodvar,
  referencevar
)
```

Arguments

data	data in long format
covar	covariates and base depvar complete
depvar	dependent variable
treatvar	treatment group
idvar	patient id
timevar	time variable for repeated visit
M	number imputations
reference	reference group must be NULL
method	RBI method must be NULL
methodvar	column location in data specifying individual RBI methods
referencevar	column location in data specifying individual reference group for RBI method

Details

checks methodvar finds missingness pattern

Value

list of outputs

Index