Package 'mimix'

February 23, 2021

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

Version 0.0.10

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

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

Description An implementation of the mimix Stata package written by Suzie Cro.
Imputes missing numerical outcomes for a longitudinal trial with protocol deviation
under distinct treatment arm-based assumptions for the unobserved data,
following the general algorithm of Carpenter, Roger, and Kenward (2013).
The imputation methods are missing at random (mar),jump to reference (j2r),
last mean carried forward (lmcf),copy increments in reference (cir) and copy refrence (cr).
Additonal options imple-
ment the Causal model (see White, Joseph and Best, 2019) and the Delta method.
The j2r,cir and cr methods must specify the reference treatment arm.
Individual-
specific imputation methods (and their reference groups) can be specified by including said
values as columns in the input data set. Both methods must not be simultaneously specified.
The M imputed data sets are output concatenated as one large data-set appended to
the original unimputed data-set. Conversion to mids data type using the mice package enables
modelling under Rubin's rules.

URL https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#sensitivity-analysis,

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` *acupuncture.*

Description

A dataset containing ~~trial~~ results of a randomised, double-blind, parallel-group ~~studies~~ comparing active treatment with placebo

Usage

```
acupuncture
```

Format

An object of class `data.frame` with 802 rows and 11 columns.

~~Details~~

@format A data frame with 802 rows and 11 columns

id

time

age

sex

migraine

chronicity

practice_id

treat

head_base

head

withdrawal_reason

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` *~~antidepressant.~~ A data set containing antidepressant trial data as described in paper by White,Royes,Best*

Description

antidepressant. A data set containing antidepressant trial data as described in paper by White,Royes,Best

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
change
miss_flag
methodcol
referencecol

`asthma` *asthma.*

Description

A dataset containing asthma trial data as used in the Stata mimix help file

Usage

```
asthma
```

Format

An object of class `data.frame` with 732 rows and 5 columns.

Details

@format A data frame containing 732 rows and 5 columns

id
time
treat
base
fev

`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 distinguishing from post-discontinuation

Usage

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

Arguments

impdata	the data with missing values
interims	the interim cases with estimated MAR values
Mimp	the number of imputations specified , ie M total imputsations

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

Arguments

`data`	Dataset in wide format
`covar`	Covariates ~~may include the~~ baseline ~~value of depvar~~. Must be complete (no missing values).
`depvar`	Dependent (outcome) variable
`treatvar`	Treatment group, coded 1,2,..
`idvar`	Participant id
`timevar`	Time point for repeated measure
`M`	Number of imputations to be created
`reference`	Reference group for J2R, CIR, CR methods
`method`	Reference-based imputation method: must be ...
`seed`	Seed value. Specify this so that a new run of the command will give the same imputed values.

prior	Prior for the variance-covariance matrix when fitting multivariate normal distributions. 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.
methodvar	vector designating variable in data-set specifying individual method
referencevar	vector designating variable in data-set specifying reference group corresponding to individual method
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
mle	logical option to Use maximum likelihood parameter estimates instead of MCMC draw parameters

Details

This is based on Suzie Cro's Stata program

sets up a 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

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

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

Examples

```
## Not run:
mimixout<-mimix("asthma",c("base"),"fev","treat","id","time",5,1,"J2R",,,,,,,,,1,0.5,)
library(mice)
fit<-with(data= as.mids(mimixout),expr = lm(fev.12~treat+base))
summary(pool(fit))
mimix("acupuncture",c("head_base"),"head","treat","id","time",1000,1,"CIR",54321,
    "jeffreys",1000,NULL,NULL,NULL,NULL,NULL,K0=1,K1=1,mle=0  )

## End(Not run)
```

mimix_package *mimix: A package porting the Stata mimix command*

Description

The mimix package provides the functionality of the Stata package plus delta and causal methods

Comparison with Stata

The mimix functions add the causal method and delta extra to the Stata version ...

Comparison with SAS

The mimix functions ... The mimix functions causal method and delta extra to the Stata version ... are similar to the SAS functions found in ..

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_

## 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