Package 'mimixR'

March 26, 2021

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

Version 0.0.14

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://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 outcome variable

withdrawal_reason

Examples

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

## 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 outcome variable
change
miss_flag
methodcol individual-specific method
referencecol individual-specific reference arm

Examples

```
## Not run:
 # Run with  covariates "basval" and "PATIENT.SEX" using columns within data to specify
 # method and reference indivdually specified columns
 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(subset(impIndiv,VISIT.NUMBER==7)),
               lm(HAMD17.TOTAL~TREATMENT.NAME+basval+PATIENT.SEX))
 summary(pool(fit))
 impantdep <- mimix(data=antidepressant,covar=c(basval,PATIENT.SEX),depvar=HAMD17.TOTAL,
         treatvar=TREATMENT.NAME,idvar=PATIENT.NUMBER,
         timevar=VISIT.NUMBER,method="J2R",reference=1,M=5,seed=54321)
 fitdep21<-with(data= as.mids(subset(impantdep,VISIT.NUMBER==7)),
              lm(HAMD17.TOTAL~TREATMENT.NAME+basval))
 summary(pool(fitdep21))

## 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 outcome variable

Examples

```
## Not run:
 impJ2R<-(mimix(data=asthma,covar=c(base),depvar=fev,treatvar=treat,idvar=id,timevar=time,
    method="J2R",reference=3,M=5,seed=101,burnin=1000)
 library(mice)
 fitJ2R<-with(data=as.mids(subset(impJ2R,time==12)),lm(fev~treat+base))
 summary(pool(fitJ2R))
 # recode treatment from numeric 2,3 to character
 # asthma$treat<- ifelse(asthma$treat==2,"placebo","active")
 # reference arm placebo
 impJ2Rridge<-(mimix(data=asthma,covar=c(base),depvar=fev,treatvar=treat,idvar=id,timevar=time,
     method="J2R",reference="placebo",delta=c(0.5,0.5,1,1 ),M=5,seed=101,prior="ridge")
 fitJ2Rridge<-with(data=as.mids(subset(impJ2Rridge,time==12)),lm(fev~treat+base))
 summary(pool(fitJ2Rridge))

## End(Not run)
```

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

Description

main wrapper for running mimix

Usage

```
mimix(
  data,
  covar = NULL,
  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	Outcome variable
treatvar	Treatment group, can be numeric or character
idvar	Participant identifier.
timevar	Time point for repeated measure
method	Reference-based imputation method: must be "J2R", "CR", "CIR","MAR", "Causal" or "LMCF"
reference	Reference group for "J2R", "CIR", "CR" methods , can be numeric or string
methodvar	column in data-set specifying individual method
referencevar	column in data-set specifying reference group as for individual method,
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 option - not recommended !

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.

Value

The M imputed data sets are output concatenated as one large dataframe in long format

appended to the original unimputed data-set

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(subset(mimixout[[2]],time==12)), lm(fev~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 based on the Stata version and similar is available in the Five_macros suite of SAS (also relevant are the mymcmc and RM_Conj SAS macros)

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

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

User_guide_to_5macros_38.pdf Roger J. (2017)

Index