# Package ‘MIIPW’

April 21, 2025

**Type** Package

**Title** IPW and Mean Score Methods for Time-Course Missing Data

**Version** 0.1.2

**Maintainer** Atanu Bhattacharjee <`atanustat@gmail.com`>

**Description** Contains functions for data analysis of Repeated measurement using GEE. Data may contain missing value in response and covariates. For parameter estimation through Fisher Scoring algorithm, Mean Score and Inverse Probability Weighted method combining with Multiple Imputation are used when there is missing value in covariates/response.
Reference for mean score method, inverse probability weighted method is
Wang et al(2007)<doi:10.1093/biostatistics/kxl024>.

**Imports** mice,Matrix,MASS

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**Depends** R (>= 2.10), lme4, spatstat

**RoxygenNote** 7.3.2

**NeedsCompilation** no

**Author** Atanu Bhattacharjee [aut, cre, ctb],
Bhrigu Kumar Rajbongshi [aut, ctb],
Gajendra K Vishwakarma [aut, ctb]

**Repository** CRAN

**Suggests** knitr, rmarkdown

**VignetteBuilder** knitr

**Date/Publication** 2025-04-21 19:10:02 UTC

## Contents

AIPW . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
lmeaipw . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
lmeipw . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

lmemeanscore . . . 11
MeanScore . . . 14
miAIPW . . . 16
miSIPW . . . 19
print_ipw . . . 21
print_meanscore . . . 22
QICmiipw . . . 22
SIPW . . . 23
srdata1 . . . 26
summary_ipw . . . 26
summary_meanscore . . . 27
updateALpha . . . 27
updateBeta . . . 28
UpdatePhi . . . 29
updateSandW . . . 29

**Index** **31**

---

`AIPW` *Fit a geeglm model using AIPW*

---

## Description

provides augmented inverse probability weighted estimates of parameters for GEE model of response variable using different covariance structure

## Usage

```
AIPW(
  data,
  formula,
  id,
  visit,
  family,
  init.beta = NULL,
  init.alpha = NULL,
  init.phi = NULL,
  tol = 0.001,
  weights = NULL,
  corstr = "independent",
  maxit = 50,
  m = 2,
  pMat,
  method = NULL
)
```

## Arguments

| | |
|---|---|
| `data` | longitudinal data set where each subject's outcome has been measured at same time points and number of visits for each patient is similar. Covariance structure of the outcome variable like "unstructured","independent","AR-1" ,"exchangeable" |
| `formula` | formula for the response model |
| `id` | column name of id of subjects in the dataset |
| `visit` | column name of timepoints of visit in the dataset |
| `family` | name of the distribution for the response variable, For more information on how to use `family` objects, see `family` |
| `init.beta` | initial values for the regression coefficient of GEE model |
| `init.alpha` | initial values for the correlation structure |
| `init.phi` | initial values for the csale parameter for |
| `tol` | tolerance in calculation of coefficients |
| `weights` | A vector of weights for each observation. If an observation has weight 0, it is excluded from the calculations of any parameters. Observations with a NA anywhere (even in variables not included in the model) will be assigned a weight of 0. Weights are updated as the mentioned the details. |
| `corstr` | a character string specifying the correlation structure. It could "independence", "exchangeable", "AR-1", "unstructured" |
| `maxit` | maximum number iteration for newton-raphson |
| `m` | number of imputation used to update the missing score function value due incomplete data. |
| `pMat` | predictor matrix as obtained in `mice` |
| `method` | method option for mice model,for information see mice |

## Details

AIPW

It uses the inverse probability weighted method to reduce the bias due to missing values in GEE model for longitudinal data. The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(\frac{\delta_{ij}}{\pi_{ij}}S(Y_{ij},\mathbf{X}_{ij},\mathbf{X}'_{ij}) + (1-\frac{\delta_{ij}}{\pi_{ij}})\phi(\mathbf{V}=\mathbf{v})) = 0$$

where $\delta_{ij} = 1$ if there is missing value in covariates and 0 otherwise, $\mathbf{X}$ is fully observed all subjects and $\mathbf{X}'$ is partially missing, where $\mathbf{V} = (Y, \mathbf{X})$. The missing score function values due to incomplete data are estimated using an imputation model through mice which we have considered as $\phi(\mathbf{V} = \mathbf{v})$).

## Value

A list of objects containing the following objects

**call** details about arguments passed in the function

**beta** estimated regression coeffictient value for the response model

**niter** number of iteration required

**betalist** list of beta values at different iteration

**weight** estimated weights for the observations

**mu** mu values according glm

**phi** etsimated phi value for the `glm` model

**hessian** estimated hessian matrix obtained from the last iteration

**betaSand** sandwich estimator value for the variance covariance matrix of the beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

SIPW,miSIPW,miAIPW

## Examples

```
 ## Not run:
##
formula<-C6kine~ActivinRIB+ActivinRIIA+ActivinRIIAB+Adiponectin+AgRP+ALCAM
pMat<-mice::make.predictorMatrix(srdata1[names(srdata1)%in%all.vars(formula)])
m1<-AIPW(data=srdata1,
formula<-formula,id='ID',
visit='Visit',family='gaussian',init.beta = NULL,
init.alpha=NULL,init.phi=1,tol=.00001,weights = NULL,
corstr = 'exchangeable',maxit=50,m=3,pMat=pMat)
##

## End(Not run)
```

---

`lmeaipw` *Fits a marginal model using AIPW*

---

## Description

provides augmented inverse probability weighted estimates of parameters for semiparametric marginal model of response variable of interest. The augmented terms are estimated by using multiple imputation model.

## Usage

```
lmeaipw(
  data,
  M = 5,
  id,
  analysis.model,
  wgt.model,
  imp.model,
  qpoints = 4,
  psiCov,
  nu,
  psi,
  sigma = NULL,
  sigmaMiss,
  sigmaR,
  dist,
  link,
  conv = 1e-04,
  maxiter,
  maxpiinv = -1,
  se = TRUE,
  verbose = FALSE
)
```

## Arguments

| | |
|---|---|
| `data` | longitudinal data with each subject specified discretely |
| `M` | number of imputation to be used in the estimation of augmentation term |
| `id` | cloumn names which shows identification number for each subject |
| `analysis.model` | A formula to be used as analysis model |
| `wgt.model` | Formula for weight model, which consider subject specific random intercept |
| `imp.model` | For for missing response imputation, which consider subject specific random intercept |
| `qpoints` | Number of quadrature points to be used while evaluating the numerical integration |

| | |
|---|---|
| `psiCov` | working model parameter |
| `nu` | working model parameter |
| `psi` | working model parameter |
| `sigma` | working model parameter |
| `sigmaMiss` | working model parameter |
| `sigmaR` | working model parameter |
| `dist` | distribution for imputation model. Currently available options are Gaussian, Binomial |
| `link` | Link function for the mean |
| `conv` | convergence tolerance |
| `maxiter` | maximum number of iteration |
| `maxpiinv` | maximum value pi can take |
| `se` | Logical for Asymptotic SE for regression coefficient of the regression model. |
| `verbose` | logical argument |

## Details

lmeaipw

It uses the augmented inverse probability weighted method to reduce the bias due to missing values in response model for longitudinal data. The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{n}\sum_{j=t_1}^{t_k}\int_{a_i}\int_{b_i}(\frac{\delta_{ij}}{\hat{\pi}_{ij}(a_i)}S(Y_{ij},\mathbf{X}_{ij};\beta)+(1-\frac{\delta_{ij}}{\hat{\pi}_{ij}(a_i)})\phi(\mathbf{V}_{ij},b_i;\psi))da_idb_i = 0$$

where $\delta_{ij} = 1$ if there is missing value in the response and 0 otherwise, $\mathbf{X}$ is fully observed all subjects, where $\mathbf{V}_{ij} = (\mathbf{X}_{ij}, A_{ij})$. The missing score function values due to incomplete data are estimated using an imputation model through FCS (here we have considered a mixed effect model) which we have considered as $\phi(\mathbf{V}_{ij} = \mathbf{v}_{ij})$). The estimated value $\hat{\phi}(\mathbf{V}_{ij} = \mathbf{v}_{ij})$) is obtained through multiple imputation. The working model for imputation of missing response is

$$Y_{ij}|b_i \sim N(\mathbf{V}_{ij}\psi + b_i, \sigma)\ ; b_i \sim N(0, \sigma_{miss})$$

and for the missing data probability

$$Logit(P(\delta_{ij} = 1|\mathbf{V}_{ij}\nu + a_i))\ ; a_i \sim N(0, \sigma_R)$$

## Value

A list of objects containing the following objects

**Call** details about arguments passed in the function

**nr.conv** logical for checking convergence in Newton Raphson algorithm

**nr.iter** number of iteration required

**nr.diff** absolute difference for roots of Newton Raphson algorithm

**beta** estimated regression coefficient for the analysis model

**var.beta** Asymptotic SE for beta

### Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

### References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

### See Also

SIPW,miSIPW,miAIPW

### Examples

```
 ## Not run:
##
library(JMbayes2)
library(lme4)
library(insight)
library(numDeriv)
library(stats)
lmer(log(alkaline)~drug+age+year+(1|id),data=na.omit(pbc2))
data1<-pbc2
data1$alkaline<-log(data1$alkaline)
names(pbc2)
apply(pbc2,2,function(x){sum(is.na(x))})
r.ij<-ifelse(is.na(data1$alkaline)==T,0,1)
data1<-cbind.data.frame(data1,r.ij)
data1$drug<-factor(data1$drug,levels=c("placebo","D-penicil"),labels = c(0,1))
data1$sex<-factor(data1$sex,levels=c('male','female'),labels=c(1,0))
data1$drug<-as.numeric(as.character(data1$drug))
data1$sex<-as.numeric(as.character(data1$sex))
r.ij~year+age+sex+drug+serBilir+(1|id)
model.r<-glmer(r.ij~year+age+sex+drug+serBilir+(1|id),family=binomial(link='logit'),data=data1)
model.y<-lmer(alkaline~year+age+sex+drug+serBilir+(1|id),data=na.omit(data1))
nu<-model.r@beta
psi<-model.y@beta
sigma<-get_variance_residual(model.y)
sigmaR<-get_variance(model.r)$var.random
sigmaMiss<-get_variance(model.y)$var.random
m11<-lmeaipw(data=data1,id='id',
analysis.model = alkaline~year,
wgt.model=~year+age+sex+drug+serBilir+(1|id),
```

```
imp.model = ~year+age+sex+drug+serBilir+(1|id),
psiCov = vcov(model.y),nu=nu,psi=psi,
sigma=sigma,sigmaMiss=sigmaMiss,sigmaR=sigmaR,dist='gaussian',link='identity',
maxiter = 200)
m11
##

## End(Not run)
```

---

## `lmeipw` *Fits a marginal model using IPW*

### Description

provides inverse probability weighted estimates of parameters for semiparametric marginal model of response variable of interest. The weights are computed using a generalized linear mixed effect model.

### Usage

```
lmeipw(
  data,
  M = 5,
  id,
  analysis.model,
  wgt.model,
  qpoints = 4,
  nu,
  sigmaR,
  dist,
  link,
  conv = 1e-04,
  maxiter,
  maxpiinv = -1,
  se = TRUE,
  verbose = FALSE
)
```

### Arguments

| | |
|---|---|
| `data` | longitudinal data with each subject specified discretely |
| `M` | number of imputation to be used in the estimation of augmentation term |
| `id` | cloumn names which shows identification number for each subject |
| `analysis.model` | A formula to be used as analysis model |
| `wgt.model` | Formula for weight model, which consider subject specific random intercept |
| `qpoints` | Number of quadrature points to be used while evaluating the numerical integration |

| | |
|---|---|
| `nu` | working model parameter |
| `sigmaR` | working model parameter |
| `dist` | distribution for imputation model. Currently available options are Gaussian, Binomial |
| `link` | Link function for the mean |
| `conv` | convergence tolerance |
| `maxiter` | maximum number of iteration |
| `maxpiinv` | maximum value pi can take |
| `se` | Logical for Asymptotic SE for regression coefficient of the regression model. |
| `verbose` | logical argument |

## Details

lmeipw

It uses the simple inverse probability weighted method to reduce the bias due to missing values in response model for longitudinal data.The response variable $\mathbf{Y}$ is related to the covariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{n}\sum_{j=t_1}^{t_k}\int_{a_i}\frac{\delta_{ij}}{\hat{\pi}_{ij}(a_i)}S(Y_{ij},\mathbf{X}_{ij})da_i = 0$$

where $\delta_{ij} = 1$ if there is missing no value in response and 0 otherwise. $\mathbf{X}$ is fully observed all subjects and for the missing data probability

$$Logit(P(\delta_{ij} = 1|\mathbf{V}_{ij}\nu + a_i))\ ; a_i \sim N(0,\sigma_R)$$

; where $\mathbf{V}_{ij} = (\mathbf{X}_{ij}, A_{ij})$

## Value

A list of objects containing the following objects

**Call** details about arguments passed in the function

**nr.conv** logical for checking convergence in Newton Raphson algorithm

**nr.iter** number of iteration required

**nr.diff** absolute difference for roots of Newton Raphson algorithm

**beta** estimated regression coefficient for the analysis model

**var.beta** Asymptotic SE for beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

SIPW,miSIPW,miAIPW

## Examples

```
 ## Not run:
##
library(JMbayes2)
library(lme4)
library(insight)
library(numDeriv)
library(stats)
lmer(log(alkaline)~drug+age+year+(1|id),data=na.omit(pbc2))
data1<-pbc2
data1$alkaline<-log(data1$alkaline)
names(pbc2)
apply(pbc2,2,function(x){sum(is.na(x))})
r.ij<-ifelse(is.na(data1$alkaline)==T,0,1)
data1<-cbind.data.frame(data1,r.ij)
data1$drug<-factor(data1$drug,levels=c("placebo","D-penicil"),labels = c(0,1))
data1$sex<-factor(data1$sex,levels=c('male','female'),labels=c(1,0))
data1$drug<-as.numeric(as.character(data1$drug))
data1$sex<-as.numeric(as.character(data1$sex))
r.ij~year+age+sex+drug+serBilir+(1|id)
model.r<-glmer(r.ij~year+age+sex+drug+serBilir+(1|id),family=binomial(link='logit'),data=data1)
nu<-model.r@beta
sigmaR<-get_variance(model.r)$var.random
m11<-lmeipw(data=data1,id='id',
            analysis.model = alkaline~year,
            wgt.model=~year+age+sex+drug+serBilir+(1|id),
            nu=nu,sigmaR=sigmaR,dist='gaussian',link='identity',qpoints=4,
            maxiter = 200)
m11
##

## End(Not run)
```

## `lmemeanscore` *Fits a marginal model using meanscore*

### Description

provides meanscore estimates of parameters for semiparametric marginal model of response variable of interest. The augmented terms are estimated by using multiple imputation model.

### Usage

```
lmemeanscore(
  data,
  M = 5,
  id,
  analysis.model,
  imp.model,
  qpoints = 4,
  psiCov,
  psi,
  sigma = NULL,
  sigmaMiss,
  dist,
  link,
  conv = 1e-04,
  maxiter,
  maxpiinv = -1,
  se = TRUE,
  verbose = FALSE
)
```

### Arguments

| | |
|---|---|
| `data` | longitudinal data with each subject specified discretely |
| `M` | number of imputation to be used in the estimation of augmentation term |
| `id` | cloumn names which shows identification number for each subject |
| `analysis.model` | A formula to be used as analysis model |
| `imp.model` | For for missing response imputation, which consider subject specific random intercept |
| `qpoints` | Number of quadrature points to be used while evaluating the numerical integration |
| `psiCov` | working model parameter |
| `psi` | working model parameter |
| `sigma` | working model parameter |
| `sigmaMiss` | working model parameter |

| | |
|---|---|
| `dist` | distribution for imputation model. Currently available options are Gaussian, Binomial |
| `link` | Link function for the mean |
| `conv` | convergence tolerance |
| `maxiter` | maximum number of iteration |
| `maxpiinv` | maximum value pi can take |
| `se` | Logical for Asymptotic SE for regression coefficient of the regression model. |
| `verbose` | logical argument |

## Details

lmemeanscore

It uses the mean score method to reduce the bias due to missing values in response model for longitudinal data.The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{n} \sum_{j=t_1}^{t_k} \int_{b_i} (\delta_{ij} S(Y_{ij}, \mathbf{X}_{ij}) + (1 - \delta_{ij}) \phi(\mathbf{V}_{ij}, b_i; \psi)) db_i = 0$$

where $\delta_{ij} = 1$ if there is missing value in response and 0 otherwise, $\mathbf{X}$ is fully observed all subjects and $\mathbf{V}_{ij} = (\mathbf{X}_{ij}, A_{ij})$. The missing score function values due to incomplete data are estimated using an imputation model through mice which we have considered as

$$Y_{ij}|b_i \sim N(\mathbf{V}_{ij}\gamma + b_i, \sigma)\ ; b_i \sim N(0, \sigma_{miss})$$

through multiple imputation.

## Value

A list of objects containing the following objects

**Call** details about arguments passed in the function

**nr.conv** logical for checking convergence in Newton Raphson algorithm

**nr.iter** number of iteration required

**nr.diff** absolute difference for roots of Newton Raphson algorithm

**beta** estimated regression coefficient for the analysis model

**var.beta** Asymptotic SE for beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

SIPW,miSIPW,miAIPW

## Examples

```
 ## Not run:
##
library(JMbayes2)
library(lme4)
library(insight)
library(numDeriv)
library(stats)
lmer(log(alkaline)~drug+age+year+(1|id),data=na.omit(pbc2))
data1<-pbc2
data1$alkaline<-log(data1$alkaline)
names(pbc2)
apply(pbc2,2,function(x){sum(is.na(x))})
r.ij<-ifelse(is.na(data1$alkaline)==T,0,1)
data1<-cbind.data.frame(data1,r.ij)
data1$drug<-factor(data1$drug,levels=c("placebo","D-penicil"),labels = c(0,1))
data1$sex<-factor(data1$sex,levels=c('male','female'),labels=c(1,0))
data1$drug<-as.numeric(as.character(data1$drug))
data1$sex<-as.numeric(as.character(data1$sex))
model.y<-lmer(alkaline~year+age+sex+drug+serBilir+(1|id),data=na.omit(data1))
psi<-model.y@beta
sigma<-get_variance_residual(model.y)
sigmaMiss<-get_variance(model.y)$var.random
m11<-lmemeanscore(data=data1,id='id',
analysis.model = alkaline~year,
imp.model = ~year+age+sex+drug+serBilir+(1|id),
psiCov = vcov(model.y),psi=psi,
sigma=sigma,sigmaMiss=sigmaMiss,dist='gaussian',link='identity',qpoints = 4,
maxiter = 200)
m11
##

## End(Not run)
```

---

## `MeanScore` *Fit a geeglm model using meanScore*

---

### Description

provides mean score estimates of parameters for GEE model of response variable using different covariance structure

### Usage

```
MeanScore(
  data,
  formula,
  id,
  visit,
  family,
  init.beta = NULL,
  init.alpha = NULL,
  init.phi = NULL,
  tol = 0.001,
  weights = NULL,
  corstr = "independent",
  maxit = 50,
  m = 2,
  pMat,
  method = NULL
)
```

### Arguments

| | |
|---|---|
| `data` | longitudinal data set where each subject's outcome has been measured at same time points and number of visits for each patient is similar. Covariance structure of the outcome variable like "unstructured","independent" ,"AR-1","exchnageable" |
| `formula` | formula for the response model |
| `id` | column name of id of subjects in the dataset |
| `visit` | column name of timepoints of visit in the dataset |
| `family` | name of the distribution for the response variable, For more information on how to use `family` objects, see `family` |
| `init.beta` | initial values for the regression coefficient of GEE model |
| `init.alpha` | initial values for the correlation structure |
| `init.phi` | initial values for the scale parameter |
| `tol` | tolerance in calculation of coefficients |

| | |
|---|---|
| `weights` | A vector of weights for each observation. If an observation has weight 0, it is excluded from the calculations of any parameters. Observations with a NA anywhere (even in variables not included in the model) will be assigned a weight of 0. Weights are updated as the mentioned the details. |
| `corstr` | a character string specifying the correlation structure. It could "independence", "exchangeable", "AR-1", "unstructured" |
| `maxit` | maximum number iteration for newton-raphson |
| `m` | number of imputation used to update the missing score function value due incomplete data. |
| `pMat` | predictor matrix as obtained in `mice` |
| `method` | method option for mice model,for information see mice |

## Details

meanScore

It uses the mean score method to reduce the bias due to missing covariate in GEE model.The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(\delta_{ij}S(Y_{ij},\mathbf{X}_{ij},\mathbf{X}'_{ij}) + (1-\delta_{ij})\phi(\mathbf{V}=\mathbf{v})) = 0$$

where $\delta_{ij} = 1$ if there is missing value in covariates and 0 otherwise, $\mathbf{X}$ is fully observed all subjects and $\mathbf{X}'$ is partially missing, where $\mathbf{V} = (Y, \mathbf{X})$. The missing score function values due to incomplete data are estimated using an imputation model through mice which we have considered as $\phi(\mathbf{V} = \mathbf{v}))$. The estimated value $\phi(\mathbf{V} = \mathbf{v}))$ is obtained through multiple imputation.

## Value

A list of objects containing the following objects

**call** details about arguments passed in the function

**beta** estimated regression coefficitient value for the response model

**niter** number of iteration required

**betalist** list of beta values at different iteration

**weight** estimated weights for the observations

**mu** mu values according glm

**phi** etsimated phi value for the `glm` model

**hessian** estimated hessian matrix obtained from the last iteration

**betaSand** sandwich estimator value for the variance covariance matrix of the beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

SIPW,miSIPW,miAIPW

## Examples

```
 ## Not run:
##
formula<-C6kine~ActivinRIB+ActivinRIIA+ActivinRIIAB+Adiponectin+AgRP+ALCAM
pMat<-mice::make.predictorMatrix(srdata1[names(srdata1)%in%all.vars(formula)])
m1<-MeanScore(data=srdata1,
formula<-formula,id='ID',
visit='Visit',family='gaussian',init.beta = NULL,
init.alpha=NULL,init.phi=1,tol=.00001,weights = NULL,
corstr = 'exchangeable',maxit=50,m=2,pMat=pMat)
##

## End(Not run)
```

---

# `miAIPW` *Fit a geeglm model using miAIPW*

## Description

provides augmented inverse probability weighted estimates of parameters for GEE model of response variable using different covariance structure. The augmented terms are estimated by using multiple imputation model.

## Usage

```
miAIPW(
  data,
  formula,
  id,
  visit,
  family,
```

```
  init.beta = NULL,
  init.alpha = NULL,
  init.phi = NULL,
  tol = 0.001,
  weights = NULL,
  corstr = "independent",
  maxit = 50,
  m = 2,
  pMat,
  method = NULL
)
```

## Arguments

| | |
|---|---|
| `data` | longitudinal data set where each subject's outcome has been measured at same time points and number of visits for each patient is similar. Covariance structure of the outcome variable like "unstuctured","independent","AR1" ,"Exchageable" |
| `formula` | formula for the response model |
| `id` | column name of id of subjects in the dataset |
| `visit` | column name of timepoints of visit in the dataset |
| `family` | name of the distribution for the response variable, For more information on how to use `family` objects, see `family` |
| `init.beta` | initial values for the regression coefficient of GEE model |
| `init.alpha` | initial values for the correlation structure |
| `init.phi` | initial values for the csale parameter for |
| `tol` | tolerance in calculation of coefficients |
| `weights` | A vector of weights for each observation. If an observation has weight 0, it is excluded from the calculations of any parameters. Observations with a NA anywhere (even in variables not included in the model) will be assigned a weight of 0. Weights are updated as the mentioned the details. |
| `corstr` | a character string specifying the correlation structure. It could "independent", "exchangeable", "AR-1", "unstructured" |
| `maxit` | maximum number iteration for newton-raphson |
| `m` | number of imputation used to update the missing score function value due incomplete data. |
| `pMat` | predictor matrix as obtained in `mice` |
| `method` | method option for mice model,for information see mice |

## Details

miAIPW

It uses the augmented inverse probability weighted method to reduce the bias due to missing values in GEE model for longitudinal data. The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) =$

$\mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(\frac{\delta_{ij}}{\pi_{ij}}S(Y_{ij},\mathbf{X}_{ij},\mathbf{X}'_{ij})+(1-\frac{\delta_{ij}}{\pi_{ij}})\phi(\mathbf{V}=\mathbf{v}))=0$$

where $\delta_{ij} = 1$ if there is missing value in covariates and 0 otherwise, $\mathbf{X}$ is fully observed all subjects and $\mathbf{X}'$ is partially missing, where $\mathbf{V} = (Y, \mathbf{X})$. The missing score function values due to incomplete data are estimated using an imputation model through mice which we have considered as $\phi(\mathbf{V} = \mathbf{v}))$. The estimated value $\phi(\mathbf{V} = \mathbf{v}))$ is obtained through multiple imputation.

## Value

A list of objects containing the following objects

**call** details about arguments passed in the function

**beta** estimated regression coeffictient value for the response model

**niter** number of iteration required

**betalist** list of beta values at different iteration

**weight** estimated weights for the observations

**mu** mu values according glm

**phi** etsimated phi value for the `glm` model

**hessian** estimated hessian matrix obtained from the last iteration

**betaSand** sandwich estimator value for the variance covariance matrix of the beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

SIPW,miSIPW,miAIPW

## Examples

```
 ## Not run:
##
formula<-C6kine~ActivinRIB+ActivinRIIA+ActivinRIIAB+Adiponectin+AgRP+ALCAM
pMat<-mice::make.predictorMatrix(srdata1[names(srdata1)%in%all.vars(formula)])
m1<-miAIPW(data=srdata1,
formula<-formula,id='ID',
 visit='Visit',family='gaussian',init.beta = NULL,
init.alpha=NULL,init.phi=1,tol=.00001,weights = NULL,
corstr = 'exchangeable',maxit=4,m=2,pMat=pMat)
##

## End(Not run)
```

---

# `miSIPW` *Fit a geeglm model using miSIPW*

## Description

provides simple inverse probability weighted estimates of parameters for GEE model of response variable using different covariance structure, missing values in covariates are multiply imputed for those subjects whose response is observed.

## Usage

```
miSIPW(
  data,
  formula,
  id,
  visit,
  family,
  init.beta = NULL,
  init.alpha = NULL,
  init.phi = NULL,
  tol = 0.001,
  weights = NULL,
  corstr = "independent",
  maxit = 50,
  m = 2,
  pMat,
  method = NULL
)
```

## Arguments

| | |
|---|---|
| `data` | longitudinal data set where each subject's outcome has been measured at same time points and number of visits for each patient is similar. Covariance structure of the outcome variable like "unstuctured","independent","AR-1" ,"exchageable" |

| | |
|---|---|
| `formula` | formula for the response model |
| `id` | column name of id of subjects in the dataset |
| `visit` | column name of timepoints of visit in the dataset |
| `family` | name of the distribution for the response variable, For more information on how to use `family` objects, see `family` |
| `init.beta` | initial values for the regression coefficient of GEE model |
| `init.alpha` | initial values for the correlation structure |
| `init.phi` | initial values for the scale parameter |
| `tol` | tolerance in calculation of coefficients |
| `weights` | A vector of weights for each observation. If an observation has weight 0, it is excluded from the calculations of any parameters. Observations with a NA anywhere (even in variables not included in the model) will be assigned a weight of 0. Weights are updated as the mentioned the details. |
| `corstr` | a character string specifying the correlation structure. It could "independence", "exchangeable", "AR-1", "unstructured" |
| `maxit` | maximum number iteration for newton-raphson |
| `m` | number of imputation used to update the missing score function value due incomplete data. |
| `pMat` | pMat predictor matrix as obtained in `mice` |
| `method` | method option for mice model,for information see mice |

## Details

miSIPW

It uses the simple inverse probability weighted method to reduce the bias due to missing values in GEE model for longitudinal data. The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{k}\sum_{j=1}^{n}\frac{\delta_{ij}}{\pi_{ij}}S(Y_{ij},\mathbf{X}_{ij},\mathbf{X}'_{ij})$$

=0 where $\delta_{ij} = 1$ if there is missing no value in covariates and 0 otherwise. $\mathbf{X}$ is fully observed all subjects and $\mathbf{X}'$ is partially missing.

## Value

A list of objects containing the following objects

**call** details about arguments passed in the function

**beta** estimated regression coeffictient value for the response model

**niter** number of iteration required

**betalist** list of beta values at different iteration

**weight** estimated weights for the observations

**mu** mu values according glm

**phi** etsimated phi value for the `glm` model

**hessian** estimated hessian matrix obtained from the last iteration

**betaSand** sandwich estimator value for the variance covariance matrix of the beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

SIPW,AIPW,miAIPW

## Examples

```
 ## Not run:
##
formula<-C6kine~ActivinRIB+ActivinRIIA+ActivinRIIAB+Adiponectin+AgRP+ALCAM
pMat<-mice::make.predictorMatrix(srdata1[names(srdata1)%in%all.vars(formula)])
m1<-miSIPW(data=srdata1,
formula=formula,id='ID',
visit='Visit',family='gaussian',init.beta = NULL,
init.alpha=NULL,init.phi=1,tol=0.001,weights = NULL,
corstr = 'exchangeable',maxit=50,m=2,pMat=pMat)
##

## End(Not run)
```

---

# `print_ipw` *print method for ipw*

## Description

print method for ipw

### Usage

```
print_ipw(x, ...)
```

### Arguments

| | |
|---|---|
| `x` | ipw object |
| `...` | further argument can be passed |

### Value

print result for ipw object

---

## `print_meanscore` *print method for meanscore*

---

### Description

print method for meanscore

### Usage

```
print_meanscore(x, ...)
```

### Arguments

| | |
|---|---|
| `x` | meanscore object |
| `...` | further argument can be passed |

### Value

print result for meanscore object

---

## `QICmiipw` *Model Selection criteria QIC*

---

### Description

It provides model selection criteria such as quasi-likelihood under the independence model criterion (QIC), an approximation to QIC under large sample i.e QICu and quasi likelihood

### Usage

```
QICmiipw(model.R, model.indep, family)
```

### Arguments

| | |
|---|---|
| `model.R` | fitted object obtained from GEE model `MeanScore,SIPW,AIPW,miSIPW,miAIPW` with correlation struture other than "independent" |
| `model.indep` | same fitted object as in `model.indep` with "independent" correlation struture |
| `family` | currently we have inlcuded "poisson","binomial","gaussian" |

### Details

QICmiipw

### Value

returns a list containing `QIC,QICu,Quasi likelihood`

### References

Pan, Wei. "Akaike's information criterion in generalized estimating equations." Biometrics 57.1 (2001): 120-125.

### Examples

```
## Not run:
 ##
 formula<-C6kine~ActivinRIB+ActivinRIIA+ActivinRIIAB+Adiponectin+AgRP+ALCAM
 pMat<-mice::make.predictorMatrix(srdata1[names(srdata1)%in%all.vars(formula)])
 m1<-MeanScore(data=srdata1,
             formula<-formula,id='ID',
             visit='Visit',family='gaussian',init.beta = NULL,
             init.alpha=NULL,init.phi=1,tol=.00001,weights = NULL,
             corstr = 'exchangeable',maxit=50,m=2,pMat=pMat)
 m11<-MeanScore(data=srdata1,
             formula<-formula,id='ID',
             visit='Visit',family='gaussian',init.beta = NULL,
             init.alpha=NULL,init.phi=1,tol=.00001,weights = NULL,
            corstr = 'independent',maxit=50,m=2,pMat=pMat)
QICmiipw(model.R=m1,model.indep=m11,family="gaussian")
##

## End(Not run)
```

---

## `SIPW` *Fit a geeglm model using SIPW*

### Description

provides simple inverse probability weighted estimates of parameters for GEE model of response variable using different covariance structure

## Usage

```
SIPW(
  data,
  formula,
  id,
  visit,
  family,
  init.beta = NULL,
  init.alpha = NULL,
  init.phi = NULL,
  tol = 0.001,
  weights = NULL,
  corstr = "independent",
  maxit = 10,
  maxvisit = NULL
)
```

## Arguments

| | |
|---|---|
| `data` | longitudinal data set where each subject's outcome has been measured at same time points and number of visits for each patient is similar. Covariance structure of the outcome variable like "unstructured","independent" ,"exchangeable" |
| `formula` | formula for the response model |
| `id` | column name of id of subjects in the dataset |
| `visit` | column name of timepoints of visit in the dataset |
| `family` | name of the distribution for the response variable, For more information on how to use `family` objects, see `family` |
| `init.beta` | initial values for the regression coefficient of GEE model |
| `init.alpha` | initial values for the correlation structure |
| `init.phi` | initial values for the scale parameter |
| `tol` | tolerance in calculation of coefficients |
| `weights` | A vector of weights for each observation. If an observation has weight 0, it is excluded from the calculations of any parameters. Observations with a NA anywhere (even in variables not included in the model) will be assigned a weight of 0. Weights are updated as the mentioned the details. |
| `corstr` | a character string specifying the correlation structure. It could "independence", "exchangeable", "AR-1", "unstructured" |
| `maxit` | maximum number of iteration |
| `maxvisit` | maximum number of visit |

## Details

SIPW

It uses the simple inverse probability weighted method to reduce the bias due to missing values in GEE model for longitudinal data.The response variable $\mathbf{Y}$ is related to the coariates as $g(\mu) = \mathbf{X}\beta$, where g is the link function for the glm. The estimating equation is

$$\sum_{i=1}^{k}\sum_{j=1}^{n}\frac{\delta_{ij}}{\pi_{ij}}S(Y_{ij},\mathbf{X}_{ij},\mathbf{X}'_{ij})$$

=0 where $\delta_{ij} = 1$ if there is missing no value in covariates and 0 otherwise. $\mathbf{X}$ is fully observed all subjects and $\mathbf{X}'$ is partially missing.

## Value

A list of objects containing the following objects

**call** details about arguments passed in the function

**beta** estimated regression coeffictient value for the response model

**niter** number of iteration required

**betalist** list of beta values at different iteration

**weight** estimated weights for the observations

**mu** mu values according glm

**phi** etsimated phi value for the `glm` model

**hessian** estimated hessian matrix obtained from the last iteration

**betaSand** sandwich estimator value for the variance covariance matrix of the beta

## Author(s)

Atanu Bhattacharjee, Bhrigu Kumar Rajbongshi and Gajendra Kumar Vishwakarma

## References

Wang, C. Y., Shen-Ming Lee, and Edward C. Chao. "Numerical equivalence of imputing scores and weighted estimators in regression analysis with missing covariates." Biostatistics 8.2 (2007): 468-473.

Seaman, Shaun R., and Stijn Vansteelandt. "Introduction to double robust methods for incomplete data." Statistical science: a review journal of the Institute of Mathematical Statistics 33.2 (2018): 184.

Vansteelandt, Stijn, James Carpenter, and Michael G. Kenward. "Analysis of incomplete data using inverse probability weighting and doubly robust estimators." Methodology: European Journal of Research Methods for the Behavioral and Social Sciences 6.1 (2010): 37.

## See Also

AIPW,miSIPW,miAIPW

### Examples

```
 ## Not run:
##
formula<-C6kine~ActivinRIB+ActivinRIIA+ActivinRIIAB+Adiponectin+AgRP+ALCAM
m1<-SIPW(data=srdata1,formula<-formula,id='ID',
visit='Visit',family='gaussian',corstr = 'exchangeable',maxit=5)
##

## End(Not run)
```

---

## `srdata1` *protein data*

### Description

Repeated measurement dataset, for each id we have four visit observations

### Usage

```
data(srdata1)
```

### Format

A dataframe with 164 rows and 9 columns

**ID** ID of subjects

**Visit** Number of times observations recorded

**C6kine,.....,GFRalpha4** These are covariates

### Examples

```
data(srdata1)
```

---

## `summary_ipw` *summary method for ipw*

### Description

summary method for ipw

### Usage

```
summary_ipw(object, ...)
```

### Arguments

| | |
|---|---|
| `object` | ipw object |
| `...` | further argument can be passed |

### Value

summary of ipw object

---

## `summary_meanscore` *summary method for meanscore*

### Description

summary method for meanscore

### Usage

```
summary_meanscore(object, ...)
```

### Arguments

| | |
|---|---|
| `object` | meanscore object |
| `...` | further argument can be passed |

### Value

summary of meanscore object

---

## `updateALpha` *internal function for updating alpha*

### Description

internal function for updating alpha

### Usage

```
updateALpha(y, x, vfun, mu, w, phi, corstr, ni, mv = NULL, id, visit)
```

### Arguments

| | |
|---|---|
| `y` | response value for GEE model |
| `x` | model matrix for the GEE model |
| `vfun` | variance function for the GLM |
| `mu` | mu vector for the GLM |
| `w` | weight matrix |
| `phi` | scale parameter |
| `corstr` | correlation structure |
| `ni` | list of visits per subject |
| `mv` | NULL |
| `id` | id column |
| `visit` | visit column |

### Details

arguments are from Fisher Scoring Algorithm

---

## `updateBeta` *internal function for updating beta through Fisher Scoring*

### Description

internal function for updating beta through Fisher Scoring

### Usage

```
updateBeta(y, x, vfun, mu, w, D, Ralpha, beta)
```

### Arguments

| | |
|---|---|
| `y` | response value for GEE model |
| `x` | model matrix for the GEE model |
| `vfun` | variance function for the GLM |
| `mu` | mu vector for the GLM |
| `w` | weight matrix |
| `D` | derivation of the inverse link function |
| `Ralpha` | correlation matrix |
| `beta` | vector of beta value for GEE model |

---

## `UpdatePhi` *internal function for updating scale parameter*

---

### Description

internal function for updating scale parameter

### Usage

```
UpdatePhi(y, x, vfun, mu, w)
```

### Arguments

| | |
|---|---|
| `y` | response value for GEE model |
| `x` | model matrix for the GEE model |
| `vfun` | variance function for the GLM |
| `mu` | mu vector for the GLM |
| `w` | weight matrix |

---

## `updateSandW` *internal function for sandwich estimator*

---

### Description

internal function for sandwich estimator

### Usage

```
updateSandW(y, x, vfun, mu, w, D, Ralpha, beta, hessmat, blockdiag)
```

### Arguments

| | |
|---|---|
| `y` | response value for GEE model |
| `x` | model matrix for the GEE model |
| `vfun` | variance function for the GLM |
| `mu` | mu vector for the GLM |
| `w` | weight matrix |
| `D` | derivation of the inverse link function |
| `Ralpha` | correlation matrix |
| `beta` | vector of beta value for GEE model |
| `hessmat` | hessian matrix |
| `blockdiag` | vector containing the dim of block matrix for block diagonal matrix |

## Details

arguments are required for obtaining Sandwich Estimator for variance matrix of regression coefficient of GEE model

# Index

∗ **datasets**
    `srdata1`, 26

`AIPW`, 2, *21*, *25*

`family`, *3*, *14*, *17*, *20*, *24*

`glm`, *4*, *15*, *18*, *21*, *25*

`lmeaipw`, 5
`lmeipw`, 8
`lmemeanscore`, 11

`MeanScore`, 14
`miAIPW`, *4*, *7*, *10*, *13*, *16*, 16, *18*, *21*, *25*
`mice`, *3*, *15*, *17*, *20*
`miSIPW`, *4*, *7*, *10*, *13*, *16*, *18*, 19, *25*

`print_ipw`, 21
`print_meanscore`, 22

`QICmiipw`, 22

`SIPW`, *4*, *7*, *10*, *13*, *16*, *18*, *21*, 23
`srdata1`, 26
`summary_ipw`, 26
`summary_meanscore`, 27

`updateALpha`, 27
`updateBeta`, 28
`UpdatePhi`, 29
`updateSandW`, 29