Package ‘mecor’

October 13, 2022

Type Package

Title Measurement Error Correction in Linear Models with a Continuous Outcome

Version 1.0.0

Author Linda Nab

Maintainer Linda Nab <lindanab4@gmail.com>


Depends R (>= 2.10)

License GPL-3

Encoding UTF-8

LazyData true

URL https://github.com/LindaNab/mecor

Imports lme4, lmerTest, numDeriv

Suggests knitr, rmarkdown

VignetteBuilder knitr

RoxygenNote 7.1.2

NeedsCompilation no

Repository CRAN

Date/Publication 2021-12-01 21:30:02 UTC
**bloodpressure**

**R topics documented:**

- bloodpressure ..................................................... 2
- haemoglobin .......................................................... 3
- haemoglobin_ext ..................................................... 4
- ipwm ................................................................. 5
- MeasError ............................................................ 8
- MeasErrorExt ......................................................... 9
- MeasErrorRandom .................................................... 10
- mecor ............................................................... 11
- sim ................................................................. 13
- sodium .............................................................. 14
- summary.mecor ....................................................... 15
- vat ................................................................. 16
- vat_ext ............................................................. 17

**Index**

<table>
<thead>
<tr>
<th>bloodpressure</th>
<th>PDAC blood pressure data [replicates study]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description**

Blood pressure, age and creatinine levels of 450 pregnant women from the Pregnancy Day Assessment Clinic.

**Usage**

bloodpressure

**Format**

A data frame with 450 rows and 6 variables:

- **creatinine** Serum creatinine (umol/L)
- **age** Age (years)
- **sbp30** Systolic blood pressure at 30 minutes (mm Hg)
- **sbp60** Systolic blood pressure at 60 minutes (mm Hg)
- **sbp90** Systolic blood pressure at 90 minutes (mm Hg)
- **sbp120** Systolic blood pressure at 120 minutes (mm Hg)

**Details**

This is a simulated dataset inspired by data that was originally published at the Dryad Digital Repository: <doi:10.5061/dryad.0bq15>
References


Examples

data("bloodpressure", package = "mecor")

---

haemoglobin

**Low-dose iron supplements haemoglobin data [internal outcome-validation study]**

Description

Capillary haemoglobin and venous haemoglobin levels of 400 subjects of a trial investigating the efficacy of low-dose iron supplements during pregnancy. Venous haemoglobin levels were observed of approximately 25% of the subjects included in the trial.

Usage

haemoglobin

Format

A data frame with 400 rows and 3 variables:

- **capillary**  Haemoglobin levels measured in capillary blood (g/L)
- **supplement**  Low-dose iron supplement (20 mg/d) (0 = no, 1 = yes)
- **venous**  Haemoglobin levels measured in venous blood (g/L)

Details

This is a simulated data set inspired by a trial investigating low-dose iron supplements <doi:10.1093/ajcn/78.1.145>. A motivating example using the example data can be found here: <doi:10.1002/sim.8359>

References


Examples

data("haemoglobin", package = "mecor")
Description

Capillary haemoglobin and venous haemoglobin levels of 100 individuals.

Usage

haemoglobin_ext

Format

A data frame with 100 rows and 2 variables:

- **capillary** Haemoglobin levels measured in capillary blood (g/L)
- **venous** Haemoglobin levels measured in venous blood (g/L)

Details

This is a simulated data set accompanying the dataset "haemoglobin", that is inspired by a trial investigating low-dose iron supplements <doi:10.1093/ajcn/78.1.145>. A motivating example using the example data can be found here: <doi:10.1002/sim.8359>

References


Examples

data("haemoglobin_ext", package = "mecor")
ipwm

Weighting for Confounding and Joint Misclassification of Exposure and Outcome

Description

ipwm implements a method for estimating the marginal causal odds ratio by constructing weights (modified inverse probability weights) that address both confounding and joint misclassification of exposure and outcome.

Usage

ipwm(
  formulas,
  data,
  outcome_true,
  outcome_mis = NULL,
  exposure_true,
  exposure_mis = NULL,
  nboot = 1000,
  conf_level = 0.95,
  fix_nNAs = FALSE,
  semiparametric = FALSE,
  optim_args = list(method = "BFGS"),
  force_optim = FALSE,
  sp = Inf,
  print = TRUE
)

Arguments

formulas a list of objects of class formula specifying the probability models for the stats:terms of some factorisation of the joint conditional probability function of exposure_true, exposure_mis, outcome_true and outcome_mis, given covariates
data data.frame containing exposure_true, exposure_mis, outcome_true, outcome_mis and covariates. Missings (NAs) are allowed on variables exposure_true and outcome_true.outcome_true a character string specifying the name of the true outcome variable that is free of misclassification but possibly unknown (NA) for some (but not all) subjectsoutcome_mis a character string specifying the name of the counterpart of outcome_true that is available on all subjects but potentially misclassifies subjects’ outcomes. The default (outcome_mis = NULL) indicates absence of outcome misclassificationexposure_true a character string specifying the name of the true exposure variable that is free of misclassification but possibly unknown (NA) for some (but not all) subjects
exposure_mis a character string specifying the name of the counterpart of exposure_true that is available on all subjects but potentially misclassifies subjects as exposed or as non-exposed. The default (exposure_mis = NULL) indicates absence of exposure misclassification.
nboot number of bootstrap samples. Setting nboot == 0 results in point estimation only.
conf_level the desired confidence level of the confidence interval.
fix_nNAs logical indicator specifying whether or not to fix the joint distribution of is.na(exposure_true) and is.na(outcome_true). If TRUE, stratified bootstrap sampling is done according to the missing data pattern.
semiparametric logical indicator specifying whether or not to parametrically sample exposure_true, exposure_mis, outcome_true and outcome_mis. If semiparametric == TRUE, it is assumed that the missing data pattern is conditionally independent of these variables given covariates. Provided nboot > 0, the missing data pattern and covariates are sampled nonparametrically. semiparametric is ignored if nboot == 0.
optim_args arguments passed onto optim if called. See Details below for more information.
force_optim logical indicator specifying whether or not to force the optim function to be called.
sp scalar shrinkage parameter in the interval (0, Inf). Values closer to zero result in greater shrinkage of the estimated odds ratio to unity; sp == Inf results in no shrinkage.
print logical indicator specifying whether or not to print the output.

Details
This function is an implementation of the weighting method described by Penning de Vries et al. (2018). The method defaults to the estimator proposed by Gravel and Platt (2018) in the absence of exposure misclassification.
The function assumes that the exposure or the outcome has a misclassified version. An error is issued when both outcome_mis and exposure_mis are set to NULL.
Provided force_optim = FALSE, ipwm is considerably more efficient when the optim function is not invoked; i.e., when (1) exposure_mis = NULL and the formula for outcome_true does not contain stats::terms involving outcome_mis or exposure_true, (2) outcome_mis = NULL and the formula for exposure_true does not contain stats::terms involving exposure_mis or outcome_true, or (3) all(is.na(data[, exposure_true]) == is.na(data[, outcome_true])) and the formulas for exposure_true and outcome_true do not contain stats::terms involving exposure_mis or outcome_mis. In these cases, ipwm uses iteratively reweighted least squares via the glm function for maximum likelihood estimation. In all other cases, optim_args is passed on to optim for optimisation of the joint likelihood of outcome_true, outcome_mis, exposure_true and exposure_mis.

Value
ipwm returns an object of class ipwm. The returned object is a list containing the following elements:
logOR the estimated log odds ratio;
call the matched function call.

If nboot != 0, the list also contains

SE a bootstrap estimate of the standard error for the estimator of the log odds ratio;
CI a bootstrap percentile confidence interval for the log odds ratio.

Author(s)

Bas B. L. Penning de Vries, <b.b.l.penning_de_vries@lumc.nl>

References


Examples

data(sim) # simulated data on 10 covariates, exposure A and outcome Y.
formulas <- list(
  Y ~ A + L1 + L2 + L3 + L4 + L5 + L6 + L7 + L8 + L9 + L10 + B + Z,
  A ~ L1 + L2 + L3 + L4 + L5 + L6 + L7 + L8 + L9 + L10 + B + Z,
  Z ~ L1 + L2 + L3 + L4 + L5 + L6 + L7 + L8 + L9 + L10 + B,
  B ~ L1 + L2 + L3 + L4 + L5 + L6 + L7 + L8 + L9 + L10
)
## Not run:
ipwm_out <- ipwm(
  formulas = formulas,
  data = sim,
  outcome_true = "Y",
  outcome_mis = "Z",
  exposure_true = "A",
  exposure_mis = "B",
  nboot = 200,
  sp = 1e6
)
ipwm_out

## End(Not run)
Description

This function creates a measurement error object, usually used as a covariate or the outcome in the formula argument of `mecor` if one wants to correct for the measurement error in that variable using a reference variable or a replicate measure.

Usage

`MeasError(substitute, reference, replicate, differential)`

Arguments

- `substitute`: a vector containing the error-prone measure
- `reference`: a vector containing the reference measure assumed without measurement error
- `replicate`: a vector or matrix with replicates of the error-prone measure with classical measurement error. This can either be replicates obtained by using the same measurement method as the substitute measure (replicates study) or replicates using a different measurement method than the substitute measure (calibration study).
- `differential`: a vector containing the variable to which the measurement error is differential.

Value

`MeasError` returns an object of class "MeasError".

An object of class `MeasError` is a list containing the substitute and reference (and replicate or differential if applicable) variables and has attributes `input` (the name of the substitute and reference or replicate and differential (if applicable) variables) and `call` (the matched call).

Author(s)

Linda Nab, <l.nab@lumc.nl>

Examples

```r
## measurement error in a covariate:
# internal covariate-validation study
data(vat)
with (vat, MeasError(substitute = wc,
                      reference = vat))

# replicates study
data(bloodpressure)
with (bloodpressure, MeasError(substitute = sbp30,
                                replicate = cbind(sbp60, sbp120)))

# outcome-calibration study
data(sodium)
```
MeasErrorExt

with(sodium, MeasError(substitute = recall, 
    replicate = cbind(urinary1, urinary2)))
## measurement error in the outcome:
# internal outcome-validation study
data(haemoglobin)
with(haemoglobin, MeasError(substitute = capillary, 
    reference = venous))
# internal outcome-validation study with differential measurement error in
# the dependent variable
data(haemoglobin)
with(haemoglobin, MeasError(substitute = capillary, 
    reference = venous, 
    differential = supplement))

MeasErrorExt Create an External Measurement Error Object

Description

This function creates an external measurement error object, usually used as a covariate or the outcome in the formula argument of mecor if one wants to correct for the measurement error in that variable using external data or externally estimated coefficients of the calibration model (covariate-measurement error) or measurement error model (outcome-measurement error).

Usage

MeasErrorExt(substitute, model)

Arguments

- substitute: a vector containing the error-prone measure
- model: a fitted linear model of class lm or a named list. The list contains a vector named coef: the coefficients of the calibration model or measurement error model and an optional matrix named vcov: the variance-covariance matrix of the coefficients

Value

MeasErrorExt returns an object of class "MeasErrorExt".

An object of class MeasErrorExt is a list containing the substitute variable and the fitted calibration model or measurement error model and has attributes input (the name of the substitute variable) and call (the matched call).

Author(s)

Linda Nab, <l.nab@lumc.nl>
Examples

```r
## measurement error in a outcome:
# external outcome-validation study
data(haemoglobin_ext)
# calibration model
calmod_fit <- lm(capillary ~ venous, data = haemoglobin)
# the external covariate-validation study can be used to correct for the
# measurement error in X_star in the dataset 'icvs', using the fitted
# calibration model
data(haemoglobin)
with (haemoglobin, MeasErrorExt(substitute = capillary,
                               model = calmod_fit))
# identical to:
calmod_coef <- coefficients(calmod_fit)
calmod_vcov <- vcov(calmod_fit)
with (haemoglobin, MeasErrorExt(substitute = capillary,
                                model = list(coef = calmod_coef,
                                              vcov = calmod_vcov)))
# when no external data is available, guesstimation of the coefficients of
# the calibration model can be used instead:
with (haemoglobin, MeasErrorExt(substitute = capillary,
                                model = list(coef = c('Intercept' = -7,
                                                  'venous' = 1.1))))
```

---

**MeasErrorRandom**

*Create a Random Measurement Error Object*

### Description

This function creates a random measurement error object, usually used as a covariate in the `formula` argument of `mecor` if one wants to correct for random measurement error in that variable.

### Usage

```r
MeasErrorRandom(substitute, variance)
```

### Arguments

- `substitute`: a vector containing the error-prone measure
- `variance`: a numeric quantifying the assumed variance of the random measurement error

### Value

`MeasErrorRandom` returns an object of class "MeasErrorRandom".

An object of class `MeasErrorRandom` is a list containing the substitute variable, the assumed variance of the random measurement error in that variable and, the attributes `input` (the name of the substitute variable) and `call` (the matched call).
mecor

Author(s)

Linda Nab, <l.nab@lumc.nl>

Examples

```r
## random measurement error in a covariate:
# internal covariate-validation study
data(bloodpressure)
with(bloodpressure, MeasErrorRandom(sbp30, variance = 0.25))
```

mecor: a Measurement Error Correction Package

Description

mecor provides correction methods for measurement error in a continuous covariate or outcome in linear regression models with a continuous outcome.

Usage

```r
mecor(formula, data, method = "standard", B = 0)
```

Arguments

- `formula`: an object of class `formula` (or one that is coerced to that class): a symbolic description of the regression model containing a `MeasError`, `MeasErrorExt` or `MeasErrorRandom` object in one of the covariates or the outcome.
- `data`: a data.frame, list or environment (or object coercible by `as.data.frame` to a data frame) containing the variables in the model specified in `formula`.
- `method`: a character string indicating the method used to correct for the measurement error, either "standard" (regression calibration for covariate measurement error and method of moments for outcome measurement error), "efficient" (efficient regression calibration for covariate measurement error and efficient method of moments for outcome measurement error), "valregcal" (validation regression calibration) or "mle" (maximum likelihood estimation). Defaults to "standard".
- `B`: number of bootstrap samples, defaults to 0.

Value

mecor returns an object of class "mecor". An object of class `mecor` is a list containing the following components:

- `corfit`: a list containing the corrected fit, including the coefficients of the corrected fit (`coef`) and the variance–covariance matrix of the coefficients of the corrected fit obtained by the delta method (`vcov`), and more depending on the method used.
- `uncorfit`: an `lm.fit` object of the uncorrected fit.
Author(s)

Linda Nab, <l.nab@lumc.nl>

References

L. Nab, R.H.H. Groenwold, P.M.J. Welsing, and M. van Smeden. Measurement error in continuous endpoints in randomised trials: problems and solutions

L. Nab, M. van Smeden, R.H. Keogh, and R.H.H. Groenwold. mecor: an R package for measurement error correction in linear models with continuous outcomes

Examples

```r
## measurement error in a covariate/outcome:
# internal covariate-validation study
data(vat)
out <-
mecor(ir_ln ~ MeasError(wc, reference = vat) + sex + age + tbf,
data = vat,
method = "standard",
B = 999)
# replicates study
data(bloodpressure)
mecor(creatinine ~ MeasError(sbp30, replicate = cbind(sbp60, sbp120)) + age,
data = bloodpressure,
method = "mle")
# outcome-calibration study
data(sodium)
mecor(MeasError(recall, replicate = cbind(urinary1, urinary2)) ~ diet,
data = sodium,
method = "efficient")
# external outcome-validation study
data(haemoglobin_ext)
calmod_fit <- lm(capillary ~ venous, data = haemoglobin_ext)
data(haemoglobin) # suppose reference venous is not available
mecor(MeasErrorExt(capillary, model = calmod_fit) ~ supplement,
data = haemoglobin)
# sensitivity analyses
data(vat) # suppose reference vat is not available
# guesstimate the coefficients of the calibration model:
mecor(ir_ln ~ MeasErrorExt(wc, model = list(coef = c(0.2, 0.5, -1.3, 0, 0.6))) + sex + age + tbf,
data = vat)
# assume random measurement error in wc of magnitude 0.25:
mecor(ir_ln ~ MeasErrorRandom(wc, variance = 0.25) + sex + age + tbf,
data = vat)
data(bloodpressure) # suppose replicates sbp60 and sbp60 are not available
mecor(creatinine ~ MeasErrorRandom(sbp30, variance = 25) + age,
data = bloodpressure)
```

## differential measurement error in the outcome:
# internal outcome-validation study
mecor(MeasError(capillary, reference = venous, differential = supplement) ~ supplement)
sim

data = haemoglobin,
method = "standard"

Simulated dataset for the ipwm function

Description
A simulated dataset containing 5000 observations of the covariates L1-L10, the true exposure A and true outcome Y, and the misclassified exposure B and misclassified outcome Z.

Usage
sim

Format
A data frame with 5000 rows and 14 variables:

L1 covariate, binary
L2 covariate, continuous
L3 covariate, binary
L4 covariate, continuous
L5 covariate, binary
L6 covariate, binary
L7 covariate, continuous
L8 covariate, binary
L9 covariate, binary
L10 covariate, continuous
A exposure, binary
Y outcome, binary
B misclassified exposure, binary
Z misclassified outcome, binary

Examples

data("sim", package = "mecor")
sodium

TONE sodium data [outcome-calibration study]

Description

Self-reported sodium intake and urinary sodium in the TONE study, a randomized controlled trial designed to investigate whether a reduction in sodium intake results in satisfactory blood pressure control. Two replicate urinary sodium measures were available in 50% of the subjects included in the trial.

Usage

sodium

Format

A data frame with 1000 rows and 4 variables:

- recall Sodium intake measured by a 24h recall (mg)
- diet Usual diet or sodium-lowering diet (0 = usual, 1 = sodium-lowering)
- urinary1 Sodium intake measured in urine (1st measure, mg)
- urinary2 Sodium intake measured in urine (2nd measure, mg)

Details

This is a simulated data set inspired by the TONE study <doi: 10.1016/1047-2797(94)00056-y>. A motivating example using the example data can be found here: <doi:10.1002/sim.7011>

References


Examples

data("sodium", package = "mecor")
Summarizing Measurement Error Correction

Description

summary method for class "mecor"

Usage

## S3 method for class 'mecor'
summary(object, alpha = 0.05, zerovar = FALSE, fieller = FALSE, ...)

Arguments

object 
an object of class "mecor", a result of a call to mecor.
alpha 
probability of obtaining a type II error.
zerovar 
a boolean indicating whether standard errors and confidence intervals using the zerovariance method must be added to the summary object.
fieller 
a boolean indicating whether confidence intervals using the fieller method must be added to the summary object.
... 
additional arguments affecting the summary produced

Value

The function summary.mecor returns a list of summary statistics of the fitted corrected model and fitted uncorrected model.

call 
the matched call
c 
summary of the corrected fit
uc 
summary of the uncorrected fit
B 
number of bootstrap replicates used
alpha 
alpha level used

See Also

The model fitting function mecor, summary

Examples

## measurement error in a covariate:
# internal covariate-validation study
data(vat)
mecor_fit <- mecor(ir_ln ~ MeasError(wc, reference = vat) + sex + age + tbf,
  data = vat,
  method = "standard")
summary(mecor_fit)
summary(mecor_fit, zerovar = TRUE, fieller = TRUE)
summary(mecor_fit, alpha = 0.10)

**Description**

Insulin resistance, waist circumference, sex, age, total body fat and visceral adipose tissue of 650 individuals from the Netherlands Epidemiology of Obesity (NEO) study. Visceral adipose tissue measurements were taken of approximately 40% of the individuals, at random.

**Usage**

vat

**Format**

A data frame with 650 rows and 6 variables:

- **ir_ln**: Natural logarithm of insulin resistance (fasting glucose (mmol/L) x fasting insulin (mU/L) / 22.5)
- **wc**: Waist circumference (standardised, cm)
- **sex**: Sex (0 = male, 1 = female)
- **age**: Age (years)
- **tbf**: Total body fat (standardised, %)
- **vat**: Visceral adipose tissue (standardised, cm^2)

**Details**

This is a simulated data set inspired by the NEO data <doi:10.1007/s10654-013-9801-3>. A motivating example using the example data can be found here: <doi:10.1093/aje/kwab114>

**References**


Examples
data("vat", package = "mecor")

vat_ext

Visceral adipose tissue external data [external covariate-validation study]

Description
Waist circumference, visceral adipose tissue, sex, age, and total body fat of 100 individuals

Usage
vat_ext

Format
A data frame with 100 rows and 5 variables:

- wc Waist circumference (standardised, cm)
- vat Visceral adipose tissue (standardised, cm^2)
- sex Sex (0 = male, 1 = female)
- age Age (years)
- tbf Total body fat (standardised, %)

Details
This is a simulated data set accompanying the dataset "vat", that is inspired by the NEO data <doi:10.1007/s10654-013-9801-3>. A motivating example using the example data can be found here: <doi:10.1093/aje/kwab114>

References


Examples
data("vat_ext", package = "mecor")
Index

* datasets
  bloodpressure, 2
  haemoglobin, 3
  haemoglobin_ext, 4
  sim, 13
  sodium, 14
  vat, 16
  vat_ext, 17
  bloodpressure, 2
  class, 5, 6, 8–11
  data.frame, 5
  formula, 5, 11
  glm, 6
  haemoglobin, 3
  haemoglobin_ext, 4
  ipwm, 5, 13
  list, 9
  lm, 9
  lm.fit, 11
  MeasError, 8, 11
  MeasErrorExt, 9, 11
  MeasErrorRandom, 10, 11
  mecor, 8–10, 11, 15
  optim, 6
  sim, 13
  sodium, 14
  summary, 15
  summary.mecor, 15
  vat, 16
  vat_ext, 17