Package 'msaeRB'

June 13, 2021

Type Package

Title Ratio Benchmarking for Multivariate Small Area Estimation

Version 0.2.1

Author Zenda Oka Briantiko [aut, cre],
Azka Ubaidillah [aut]

Maintainer Zenda Oka Briantiko <221710087@stis.ac.id>

Description Implements multivariate ratio benchmarking small area estimation. This package provides ratio benchmarking estimation for univariate and multivariate small area estimation and its MSE. In fact, MSE estimators for ratio benchmark are not readily available, so resampling method that called parametric bootstrap is applied. The ratio benchmark model and parametric bootstrap in this package are based on the model proposed in small area estimation. J.N.K Rao and Isabel Molina (2015, ISBN: 978-1-118-73578-7).

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

URL https://github.com/zendaokab/msaeRB

BugReports https://github.com/zendaokab/msaeRB/issues

Suggests covr, knitr, rmarkdown

VignetteBuilder knitr

Imports magic, abind, Matrix, MASS, stats

Depends R (>= 2.10)

NeedsCompilation no

Repository CRAN

Date/Publication 2021-06-12 23:50:02 UTC

R topics documented:

`datamsaeRB` *Sample Data for Multivariate Small Area Estimation with Ratio Benchmarking*

Description

Dataset to simulate ratio benchmarking of Multivariate Fay-Herriot model

This data is generated based on multivariate Fay-Herriot model by these following steps:

1. Generate explanatory variables `X1` and `X2`. `X1 ~ N(10,1)` and `X2 ~ U(9.5,10.5)`.
   Sampling error `e` is generated with the following σ_{e11} = 0.01, σ_{e22} = 0.02, σ_{e33} = 0.03, and ρ_e = 1/2.
   For random effect `u`, we set σ_{u11}= 0.02, σ_{u22}= 0.03, and σ_{u33}= 0.04.
   For the weight, we generate `w1,w2,w3` by set w1, w2, w3 ~ U(10, 20)
   Set beta, β01 = 10, β02 = 8, β03 = 6, β11 = -0.3, β12 = 0.2, β13 = 0.4, β21 = 0.5, β22 = -0.1, and β23 = -0.2.
   Calculate direct estimation `Y1 Y2 Y3` where $Y_i = X * \beta + u_i + e_i$
2. Then combine the direct estimations `Y1 Y2 Y3`, explanatory variables `X1 X2`, weight `w1 w2 w3`, and sampling varians covarians `v1 v12 v13 v2 v23 v3` in a dataframe then named as datamsaeRB

Usage

```
datamsaeRB
```

Format

A data frame with 30 rows and 14 variables:

Y1 Direct Estimation of Y1

Y2 Direct Estimation of Y2

Y3 Direct Estimation of Y3

X1 Auxiliary variable of X1

X2 Auxiliary variable of X2

w1 Known proportion of units in small areas of Y1

w2 Known proportion of units in small areas of Y2

w3 Known proportion of units in small areas of Y3

v1 Sampling Variance of Y1

v12 Sampling Covariance of Y1 and Y2

v13 Sampling Covariance of Y1 and Y3

v2 Sampling Variance of Y2

v23 Sampling Covariance of Y2 and Y3

v3 Sampling Variance of Y3

`datamsaeRBns` *Sample Data for Multivariate Non Sampled Area in Small Area Estimation with Ratio Benchmarking*

Description

Dataset to simulate ratio benchmarking of Multivariate non sampled area in Fay-Herriot model

This data is generated based on multivariate Fay-Herriot model by these following steps:

1. Generate explanatory variables `X1` and `X2`. `X1 ~ N(10,1)` and `X2 ~ U(9.5,10.5)`.
   Cluster is generated discrete uniform distribution with a = 1 and b = 2.
   Sampling error e is generated with the following σ_{e11} = 0.01, σ_{e22} = 0.02, σ_{e33} = 0.03, and ρ_e = 1/2.
   For random effect u, we set σ_{u11}= 0.02, σ_{u22}= 0.03, and σ_{u33}= 0.04.
   For the weight, we generate `w1,w2,w3` by set w1, w2, w3 ~ U(10, 20)
   Set beta, β01 = 10, β02 = 8, β03 = 6, β11 = -0.3, β12 = 0.2, β13 = 0.4, β21 = 0.5, β22 = -0.1, and β23 = -0.2.
   Calculate direct estimation `Y1 Y2 Y3` where $Y_i = X * \beta + u_i + e_i$
2. Then combine the direct estimations `Y1 Y2 Y3`, explanatory variables `X1 X2`, weight `w1 w2 w3`, and sampling varians covarians `v1 v12 v13 v2 v23 v3` in a dataframe then named as datamsaeRB

Usage

```
datamsaeRBns
```

Format

A data frame with 30 rows and 17 variables:

Y1 Direct Estimation of Y1

Y2 Direct Estimation of Y2

Y3 Direct Estimation of Y3

X1 Auxiliary variable of X1

X2 Auxiliary variable of X2

w1 Known proportion of units in small areas of Y1

w2 Known proportion of units in small areas of Y2

w3 Known proportion of units in small areas of Y3

v1 Sampling Variance of Y1

v12 Sampling Covariance of Y1 and Y2

v13 Sampling Covariance of Y1 and Y3

v2 Sampling Variance of Y2

v23 Sampling Covariance of Y2 and Y3

v3 Sampling Variance of Y3

c1 Cluster for Y1

c2 Cluster for Y2

c3 Cluster for Y3

`est_msaeRB` *EBLUPs Ratio Benchmarking based on a Multivariate Fay Herriot (Model 1)*

Description

This function gives EBLUPs ratio benchmarking based on multivariate Fay-Herriot (Model 1)

Usage

```
est_msaeRB(
  formula,
  vardir,
  weight,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

formula	an object of class list of formula describe the fitted models
vardir	matrix containing sampling variances of direct estimators. The order is: var1,cov12,...,cov1r,var2,c
weight	matrix containing proportion of units in small areas. The order is: w1,w2,...,w(r)
samevar	logical. If TRUE, the varians is same. Default is FALSE
MAXITER	maximum number of iterations for Fisher-scoring. Default is 100
PRECISION	coverage tolerance limit for the Fisher Scoring algorithm. Default value is 1e-4
data	dataframe containing the variables named in formula, vardir, and weight

Value

This function returns a list with following objects:

eblup — a list containing a value of estimators

- est.eblup : a dataframe containing EBLUP estimators
- est.eblupRB : a dataframe containing ratio benchmark estimators

fit — a list contining following objects:

- method : fitting method, named "REML"
- convergence : logical value of convergence of Fisher Scoring
- iterations : number of iterations of Fisher Scoring algorithm
- estcoef : a data frame containing estimated model coefficients (beta,std. error,t value,p-value)
- refvar : estimated random effect variance

random.effect	a data frame containing values of random effect estimators
agregation	a data frame containing agregation of direct, EBLUP, and ratio benchmark estimation

Examples

```
## load dataset
data(datamsaeRB)

# Compute EBLUP and Ratio Benchmark using auxiliary variables X1 and X2 for each dependent variable

## Using parameter 'data'
Fo = list(f1 = Y1 ~ X1 + X2,
          f2 = Y2 ~ X1 + X2,
          f3 = Y3 ~ X1 + X2)
vardir = c("v1", "v12", "v13", "v2", "v23", "v3")
weight = c("w1", "w2", "w3")

est_msae = est_msaeRB(Fo, vardir, weight, data = datamsaeRB)

## Without parameter 'data'
```

```
Fo = list(f1 = datamsaeRB$Y1 ~ datamsaeRB$X1 + datamsaeRB$X2,
          f2 = datamsaeRB$Y2 ~ datamsaeRB$X1 + datamsaeRB$X2,
          f3 = datamsaeRB$Y3 ~ datamsaeRB$X1 + datamsaeRB$X2)
vardir = datamsaeRB[, c("v1", "v12", "v13", "v2", "v23", "v3")]
weight = datamsaeRB[, c("w1", "w2", "w3")]

est_msae = est_msaeRB(Fo, vardir, weight)

## Return
est_msae$eblup$est.eblupRB # to see the Ratio Benchmark estimators
```

`est_msaeRBns` *EBLUPs Ratio Benchmarking for Non Sampled Area based on a Multivariate Fay Herriot (Model 1)*

Description

This function gives EBLUPs ratio benchmarking for non sampled area based on multivariate Fay-Herriot (Model 1)

Usage

```
est_msaeRBns(
  formula,
  vardir,
  weight,
  cluster,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

`formula`	an object of class list of formula describe the fitted models
`vardir`	matrix containing sampling variances of direct estimators. The order is: `var1,cov12,...,cov1r,var2,c`
`weight`	matrix containing proportion of units in small areas. The order is: `w1,w2,...,w(r)`
`cluster`	matrix containing cluster of auxiliary variables. The order is: `c1,c2,...,c(r)`
`samevar`	logical. If `TRUE`, the varians is same. Default is `FALSE`
`MAXITER`	maximum number of iterations for Fisher-scoring. Default is 100
`PRECISION`	coverage tolerance limit for the Fisher Scoring algorithm. Default value is `1e-4`
`data`	dataframe containing the variables named in formula, vardir, and weight

Value

This function returns a list with following objects:

`eblup` a list containing a value of estimators

- est.eblup : a dataframe containing EBLUP estimators
- est.eblupRB : a dataframe containing ratio benchmark estimators

`fit` a list contining following objects:

- method : fitting method, named "REML"
- convergence : logical value of convergence of Fisher Scoring
- iterations : number of iterations of Fisher Scoring algorithm
- estcoef : a data frame containing estimated model coefficients (`beta`,`std. error`,`t value`,`p-value`)
- refvar : estimated random effect variance

`random.effect` a data frame containing values of random effect estimators

`agregation` a data frame containing agregation of direct, EBLUP, and ratio benchmark estimation

Examples

```
## load dataset
data(datamsaeRBns)

# Compute EBLUP and Ratio Benchmark using auxiliary variables X1 and X2 for each dependent variable

## Using parameter 'data'
Fo = list(f1 = Y1 ~ X1 + X2,
          f2 = Y2 ~ X1 + X2,
          f3 = Y3 ~ X1 + X2)
vardir = c("v1", "v12", "v13", "v2", "v23", "v3")
weight = c("w1", "w2", "w3")
cluster = c("c1", "c2", "c3")

est_msae = est_msaeRBns(Fo, vardir, weight, cluster, data = datamsaeRBns)

## Without parameter 'data'
Fo = list(f1 = datamsaeRBns$Y1 ~ datamsaeRBns$X1 + datamsaeRBns$X2,
          f2 = datamsaeRBns$Y2 ~ datamsaeRBns$X1 + datamsaeRBns$X2,
          f3 = datamsaeRBns$Y3 ~ datamsaeRBns$X1 + datamsaeRBns$X2)
vardir = datamsaeRBns[, c("v1", "v12", "v13", "v2", "v23", "v3")]
weight = datamsaeRBns[, c("w1", "w2", "w3")]
cluster = datamsaeRBns[, c("c1", "c2", "c3")]

est_msae = est_msaeRBns(Fo, vardir, weight, cluster)

## Return
est_msae$eblup$est.eblupRB # to see the Ratio Benchmark estimators
```

`est_saeRB` *EBLUPs Ratio Benchmarking based on a Univariate Fay-Herriot (Model 1)*

Description

This function gives EBLUPs ratio benchmarking based on univariate Fay-Herriot (model 1)

Usage

```
est_saeRB(
  formula,
  vardir,
  weight,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

`formula`	an object of class list of formula describe the fitted model
`vardir`	vector containing sampling variances of direct estimators
`weight`	vector containing proportion of units in small areas
`samevar`	logical. If `TRUE`, the varians is same. Default is `FALSE`
`MAXITER`	maximum number of iterations for Fisher-scoring. Default is 100
`PRECISION`	coverage tolerance limit for the Fisher Scoring algorithm. Default value is `1e-4`
`data`	dataframe containing the variables named in formula, vardir, and weight

Value

This function returns a list with following objects:

`eblup` a list containing a value of estimators

- est.eblup : a dataframe containing EBLUP estimators
- est.eblupRB : a dataframe containing ratio benchmark estimators

`fit` a list contining following objects:

- method : fitting method, named "REML"
- convergence : logical value of convergence of Fisher Scoring
- iterations : number of iterations of Fisher Scoring algorithm
- estcoef : a data frame containing estimated model coefficients (`beta`, `std. error`, `t value`, `p-value`)

- refvar : estimated random effect variance

random.effect	a data frame containing values of random effect estimators
agregation	a data frame containing agregation of direct, EBLUP, and ratio benchmark estimation

Examples

```
## load dataset
data(datamsaeRB)

# Compute EBLUP and Ratio Benchmark using auxiliary variables X1 and X2 for each dependent variable

## Using parameter 'data'
est_sae = est_saeRB(Y1 ~ X1 + X2, v1, w1, data = datamsaeRB)

## Without parameter 'data'
est_sae = est_saeRB(datamsaeRB$Y1 ~ datamsaeRB$X1 + datamsaeRB$X2, datamsaeRB$v1, datamsaeRB$w1)

## Return
est_sae$eblup$est.eblupRB # to see the Ratio Benchmark estimators
```

---

est_saeRBns *EBLUPs Ratio Benchmarking for Non Sampled Area based on a Univariate Fay-Herriot (Model 1)*

Description

This function gives EBLUPs ratio benchmarking for non sampled area based on univariate Fay-Herriot (model 1)

Usage

```
est_saeRBns(
  formula,
  vardir,
  weight,
  cluster,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

formula	an object of class list of formula describe the fitted model
vardir	vector containing sampling variances of direct estimators
weight	vector containing proportion of units in small areas
cluster	vector containing cluster of auxiliary variable
samevar	logical. If TRUE, the varians is same. Default is FALSE
MAXITER	maximum number of iterations for Fisher-scoring. Default is 100
PRECISION	coverage tolerance limit for the Fisher Scoring algorithm. Default value is 1e-4
data	dataframe containing the variables named in formula, vardir, and weight

Value

This function returns a list with following objects:

eblup — a list containing a value of estimators

- est.eblup : a dataframe containing EBLUP estimators
- est.eblupRB : a dataframe containing ratio benchmark estimators

fit — a list contining following objects:

- method : fitting method, named "REML"
- convergence : logical value of convergence of Fisher Scoring
- iterations : number of iterations of Fisher Scoring algorithm
- estcoef : a data frame containing estimated model coefficients (beta, std. error, t value, p-value)
- refvar : estimated random effect variance

random.effect	a data frame containing values of random effect estimators
agregation	a data frame containing agregation of direct, EBLUP, and ratio benchmark estimation

Examples

```
## load dataset
data(datamsaeRBns)

# Compute EBLUP and Ratio Benchmark using auxiliary variables X1 and X2 for each dependent variable

## Using parameter 'data'
est_sae = est_saeRBns(Y1 ~ X1 + X2, v1, w1, c1, data = datamsaeRBns)

## Without parameter 'data'
est_sae = est_saeRBns(datamsaeRBns$Y1 ~ datamsaeRBns$X1 + datamsaeRBns$X2,
datamsaeRBns$v1, datamsaeRBns$w1, datamsaeRBns$c1)

## Return
est_sae$eblup$est.eblupRB # to see the Ratio Benchmark estimators
```

mse_msaeRB *Parametric Bootstrap Mean Squared Error Estimators of Ratio Benchmarking for Multivariate Small Area Estimation*

Description

Calculates the parametric bootstrap mean squared error estimates of ratio benchmarking for multivariate small area estimation

Usage

```
mse_msaeRB(
  formula,
  vardir,
  weight,
  samevar = FALSE,
  B = 1000,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

formula	an object of class list of formula describe the fitted models
vardir	matrix containing sampling variances of direct estimators. The order is: var1,cov12,...,cov1r,var2,c
weight	matrix containing proportion of units in small areas. The order is: w1,w2,...,w(r)
samevar	logical. If TRUE, the varians is same. Default is FALSE
B	number of bootstrap. Default is 1000
MAXITER	maximum number of iterations for Fisher-scoring. Default is 100
PRECISION	coverage tolerance limit for the Fisher Scoring algorithm. Default value is 1e-4
data	dataframe containing the variables named in formula, vardir, and weight

Value

mse.eblup	estimated mean squared errors of the EBLUPs for the small domains based on Prasad Rao
pbmse.eblupRB	parametric bootstrap mean squared error estimates of the ratio benchmark
running.time	time for running function

Examples

```
## load dataset
data(datamsaeRB)

# Compute MSE EBLUP and Ratio Benchmark
# This is the long running example
## Using parameter 'data'
Fo = list(f1 = Y1 ~ X1 + X2,
          f2 = Y2 ~ X1 + X2,
          f3 = Y3 ~ X1 + X2)
vardir = c("v1", "v12", "v13", "v2", "v23", "v3")
weight = c("w1", "w2", "w3")

mse_msae = est_msaeRB(Fo, vardir, weight, data = datamsaeRB)

## Without parameter 'data'
Fo = list(f1 = datamsaeRB$Y1 ~ datamsaeRB$X1 + datamsaeRB$X2,
          f2 = datamsaeRB$Y2 ~ datamsaeRB$X1 + datamsaeRB$X2,
          f3 = datamsaeRB$Y3 ~ datamsaeRB$X1 + datamsaeRB$X2)
vardir = datamsaeRB[, c("v1", "v12", "v13", "v2", "v23", "v3")]
weight = datamsaeRB[, c("w1", "w2", "w3")]

mse_msae = mse_msaeRB(Fo, vardir, weight)

## Return
mse_msae$pbmse.eblupRB # to see the MSE of Ratio Benchmark
```

`mse_msaeRBns` *Parametric Bootstrap Mean Squared Error Estimators of Ratio Benchmarking for Multivariate Non Sampled Area in Small Area Estimation*

Description

Calculates the parametric bootstrap mean squared error estimates of ratio benchmarking for multivariate non sampled area in small area estimation

Usage

```
mse_msaeRBns(
  formula,
  vardir,
  weight,
  cluster,
  samevar = FALSE,
  B = 1000,
  MAXITER = 100,
  PRECISION = 1e-04,
```

```
  data
)
```

Arguments

formula	an object of class list of formula describe the fitted models
vardir	matrix containing sampling variances of direct estimators. The order is: var1,cov12,...,cov1r,var2,c
weight	matrix containing proportion of units in small areas. The order is: w1,w2,...,w(r)
cluster	matrix containing cluster of auxiliary variables. The order is: c1,c2,...,c(r)
samevar	logical. If TRUE, the varians is same. Default is FALSE
B	number of bootstrap. Default is 1000
MAXITER	maximum number of iterations for Fisher-scoring. Default is 100
PRECISION	coverage tolerance limit for the Fisher Scoring algorithm. Default value is 1e-4
data	dataframe containing the variables named in formula, vardir, and weight

Value

mse.eblup	estimated mean squared errors of the EBLUPs for the small domains based on Prasad Rao
pbmse.eblupRB	parametric bootstrap mean squared error estimates of the ratio benchmark
running.time	time for running function

Examples

```
## load dataset
data(datamsaeRBns)

# Compute MSE EBLUP and Ratio Benchmark
# This is the long running example
## Using parameter 'data'
Fo = list(f1 = Y1 ~ X1 + X2,
          f2 = Y2 ~ X1 + X2,
          f3 = Y3 ~ X1 + X2)
vardir = c("v1", "v12", "v13", "v2", "v23", "v3")
weight = c("w1", "w2", "w3")
cluster = c("c1", "c2", "c3")

mse_msae = mse_msaeRBns(Fo, vardir, weight, cluster, data = datamsaeRBns)

## Without parameter 'data'
Fo = list(f1 = datamsaeRBns$Y1 ~ datamsaeRBns$X1 + datamsaeRBns$X2,
          f2 = datamsaeRBns$Y2 ~ datamsaeRBns$X1 + datamsaeRBns$X2,
          f3 = datamsaeRBns$Y3 ~ datamsaeRBns$X1 + datamsaeRBns$X2)
vardir = datamsaeRBns[, c("v1", "v12", "v13", "v2", "v23", "v3")]
weight = datamsaeRBns[, c("w1", "w2", "w3")]
cluster = datamsaeRBns[, c("c1", "c2", "c3")]
```

```
mse_msae = mse_msaeRBns(Fo, vardir, weight, cluster)

## Return
mse_msae$pbmse.eblupRB # to see the MSE of Ratio Benchmark
```

`mse_saeRB` *Parametric Bootstrap Mean Squared Error Estimators of Ratio Benchmarking for Univariate Small Area Estimation*

Description

Calculates the parametric bootstrap mean squared error estimates of ratio benchmarking for univariate small area estimation

Usage

```
mse_saeRB(
  formula,
  vardir,
  weight,
  samevar = FALSE,
  B = 1000,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

`formula`	an object of class list of formula describe the fitted model
`vardir`	vector containing sampling variances of direct estimators
`weight`	vector containing proportion of units in small areas
`samevar`	logical. If `TRUE`, the varians is same. Default is `FALSE`
`B`	number of bootstrap. Default is 1000
`MAXITER`	maximum number of iterations for Fisher-scoring. Default is 100
`PRECISION`	coverage tolerance limit for the Fisher Scoring algorithm. Default value is `1e-4`
`data`	dataframe containing the variables named in formula, vardir, and weight

Value

`mse.eblup`	estimated mean squared errors of the EBLUPs for the small domains based on Prasad Rao
`pbmse.eblupRB`	parametric bootstrap mean squared error estimates of the ratio benchmark
`running.time`	time for running function

Examples

```
## load dataset
data(datamsaeRB)

# Compute MSE EBLUP and Ratio Benchmark

## Using parameter 'data'
mse_sae = mse_saeRB(Y1 ~ X1 + X2, v1, w1, data = datamsaeRB)

## Without parameter 'data'
mse_sae = mse_saeRB(datamsaeRB$Y1 ~ datamsaeRB$X1 + datamsaeRB$X2, datamsaeRB$v1, datamsaeRB$w1)

## Return
mse_sae$pbmse.eblupRB # to see the MSE Ratio Benchmark estimators
```

---

mse_saeRBns *Parametric Bootstrap Mean Squared Error Estimators of Ratio Benchmarking for Univariate Non Sampled Area in Small Area Estimation*

Description

Calculates the parametric bootstrap mean squared error estimates of ratio benchmarking for univariate non sampled area in small area estimation

Usage

```
mse_saeRBns(
  formula,
  vardir,
  weight,
  cluster,
  samevar = FALSE,
  B = 1000,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

formula	an object of class list of formula describe the fitted model
vardir	vector containing sampling variances of direct estimators
weight	vector containing proportion of units in small areas
cluster	vector containing cluster of auxiliary variable
samevar	logical. If TRUE, the varians is same. Default is FALSE

B	number of bootstrap. Default is 1000
MAXITER	maximum number of iterations for Fisher-scoring. Default is 100
PRECISION	coverage tolerance limit for the Fisher Scoring algorithm. Default value is 1e-4
data	dataframe containing the variables named in formula, vardir, and weight

Value

mse.eblup	estimated mean squared errors of the EBLUPs for the small domains based on Prasad Rao
pbmse.eblupRB	parametric bootstrap mean squared error estimates of the ratio benchmark
running.time	time for running function

Examples

```
## load dataset
data(datamsaeRBns)

# Compute MSE EBLUP and Ratio Benchmark

## Using parameter 'data'
mse_sae = mse_saeRBns(Y1 ~ X1 + X2, v1, w1, c1, data = datamsaeRBns)

## Without parameter 'data'
mse_sae = mse_saeRBns(datamsaeRBns$Y1 ~ datamsaeRBns$X1 + datamsaeRBns$X2,
datamsaeRBns$v1, datamsaeRBns$w1, datamsaeRBns$c1)

## Return
mse_sae$pbmse.eblupRB # to see the MSE Ratio Benchmark estimators
```

Index