Package ‘msae’

April 12, 2021

Type Package

Title Multivariate Fay Herriot Models for Small Area Estimation

Version 0.1.3

Author Novia Permatasari, Azka Ubaidillah

Maintainer Novia Permatasari <16.9335@stis.ac.id>

Description Implements multivariate Fay-Herriot models for small area estimation. It uses empirical best linear unbiased prediction (EBLUP) estimator. Multivariate models consider the correlation of several target variables and borrow strength from auxiliary variables to improve the effectiveness of a domain sample size. Models which accommodated by this package are univariate model with several target variables (model 0), multivariate model (model 1), autoregressive multivariate model (model 2), and heteroscedastic autoregressive multivariate model (model 3). Functions provide EBLUP estimators and mean squared error (MSE) estimator for each model. These models were developed by Roberto Benavent and Domingo Morales (2015) <doi:10.1016/j.csda.2015.07.013>.

License GPL-2

LazyData TRUE

Encoding UTF-8

Depends R (>= 2.10)

Imports magic

RoxygenNote 7.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2021-04-12 05:10:03 UTC

R topics documented:

`datasae1` *Data generated based on Multivariate Fay Herriot Model (Model 1)*

Description

This data is generated based on multivariate Fay-Herriot model (model 1) by these following steps:

1. Generate sampling error `e`, random effect `u`, and auxiliary variables `X1` `X2`.
   - For sampling error `e`, we set $e \sim N_3(0, V_e)$, where $V_e = (\sigma_{ij})_{i,j=1,2,3}$, with σ_{11} = 0.1 , σ_{22} = 0.2 , σ_{33} = 0.3 , and ρ_e = 0.5.
   - For random effect `u`, we set $u \sim N_3(0, V_u)$, where σ_{u11} = 0.2 , σ_{u22} = 0.4 , and σ_{u33} = 1.2.
   - For auxiliary variables `X1` and `X2`, we set $X1 \sim N(5, 0.1)$ and $X2 \sim N(10, 0.2)$.
2. Calculate direct estimation `Y1` `Y2` and `Y3` , where $Y_i = X * \beta + u_i + e_i$. We take $\beta_1 = 5$ and $\beta_2 = 10$.

Auxiliary variables `X1` `X2`, direct estimation `Y1` `Y2` `Y3`, and sampling variance-covariance `v1` `v2` `v3` `v12` `v13` `v23` are combined into a dataframe called datasae1.

Usage

```
datasae1
```

Format

A data frame with 50 rows and 11 variables:

X1 Auxiliary variable of X1

X2 Auxiliary variable of X2

Y1 Direct Estimation of Y1

Y2 Direct Estimation of Y2

Y3 Direct Estimation 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

Reference

Benavent, Roberto & Morales, Domingo. (2015). Multivariate Fay-Herriot models for small area estimation. Computational Statistics & Data Analysis. 100. 372-390. DOI: 10.1016/j.csda.2015.07.013.

---

datasae2 *Data generated based on Autoregressive Multivariate Fay Herriot Model (Model 2)*

Description

This data is generated based on autoregressive multivariate Fay-Herriot model (model 2) by following these steps:

1. Generate sampling error `e`, random effect `u`, and auxiliary variables `X1 X2`.
   - For sampling error `e`, we set $e \sim N_3(0, V_e)$, where $V_e = (\sigma_{ij})_{i,j=1,2,3}$, with σ_{11} = 0.1 , σ_{22} = 0.2 , σ_{33} = 0.3 , and ρ_e = 0.5.
   - For random effect `u`, we set $u \sim N_3(0, V_u)$, where σ_u = 0.4, and ρ_u = 0.8.
   - For auxiliary variables `X1` and `X2`, we set $X1 \sim N(5, 0.1)$ and $X2 \sim N(10, 0.2)$.
2. Calculate direct estimation `Y1 Y2` and `Y3` , where $Y_i = X * \beta + u_i + e_i$. We take $\beta_1 = 5$ and $\beta_2 = 10$.

Auxiliary variables `X1 X2`, direct estimation `Y1 Y2 Y3`, and sampling variance-covariance `v1 v2 v3 v12 v13 v23` are combined into a dataframe called datasae1.

Usage

```
datasae2
```

Format

A data frame with 50 rows and 11 variables:

X1 Auxiliary variable of X1

X2 Auxiliary variable of X2

Y1 Direct Estimation of Y1

Y2 Direct Estimation of Y2

Y3 Direct Estimation 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

Reference

Benavent, Roberto & Morales, Domingo. (2015). Multivariate Fay-Herriot models for small area estimation. Computational Statistics & Data Analysis. 100. 372-390. DOI: 10.1016/j.csda.2015.07.013.

`datasae3` *Data generated based on Heteroscedastic Autoregressive Multivariate Fay Herriot Model (Model 3)*

Description

This data is generated based on heteroscedastic autoregressive multivariate Fay-Herriot model (model 3) by following these steps:

1. Generate sampling error `e`, random effect `u`, and auxiliary variables `X1 X2`.
   - For sampling error `e`, we set e ~ $N_3(0, V_e)$, where $V_e = (\sigma_{ij})_{i,j=1,2,3}$, with σ_{11} = 0.1 , σ_{22} = 0.2 , σ_{33} = 0.3 , and ρ_e = 0.5.
   - For random effect `u`, we set u ~ $N_3(0, V_u)$, where σ_{u11} = 0.2 , σ_{u22} = 0.4 , σ_{u33} = 1.2, and ρ_u = 0.8.
   - For auxiliary variables `X1 and X2`, we set $X1$ ~ $N(5, 0.1)$ and $X2$ ~ $N(10, 0.2)$.
2. Calculate direct estimation `Y1 Y2 and Y3` , where Y_i = $X * \beta + u_i + e_i$. We take $\beta_1 = 5$ and $\beta_2 = 10$.

Auxiliary variables `X1 X2`, direct estimation `Y1 Y2 Y3`, and sampling variance-covariance `v1 v2 v3 v12 v13 v23` are combined into a dataframe called datasae1.

Usage

```
datasae3
```

Format

A data frame with 50 rows and 11 variables:

X1 Auxiliary variable of X1

X2 Auxiliary variable of X2

Y1 Direct Estimation of Y1

Y2 Direct Estimation of Y2

Y3 Direct Estimation 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

Reference

Benavent, Roberto & Morales, Domingo. (2015). Multivariate Fay-Herriot models for small area estimation. Computational Statistics & Data Analysis. 100. 372-390. DOI: 10.1016/j.csda.2015.07.013.

---

df2matR *Transform Dataframe to Matrix R*

Description

This function transforms dataframe contains sampling variance to block diagonal matrix R

Usage

```
df2matR(var.df, r)
```

Arguments

`var.df`	dataframe of sampling variances of direct estimators.
`r`	number of variables

Value

Block diagonal matrix R

Examples

```
NULL
```

---

eblupMFH1 *EBLUPs based on a Multivariate Fay Herriot (Model 1)*

Description

This function gives the EBLUP and MSE based on a multivariate Fay-Herriot model (model 1)

Usage

```
eblupMFH1(
  formula,
  vardir,
  samevar = FALSE,
  MAXITER = 100,
  PRECISION = 1e-04,
  data
)
```

Arguments

`formula`	an object of class list of formula, describe the model to be fitted
`vardir`	if data is available, it is vector containing name of sampling variances of direct estimators. if not, it is data frame of sampling variances of direct estimators. The order is : `var1,var2,. ,var(k) ,cov12,. cov1k,cov23,. ,cov(k-1)(k)`
`samevar`	logical input, true if variances of the data are same, Default: `FALSE`
`MAXITER`	maximum number of iterations allowed in the Fisher-scoring algorithm, Default: `100`
`PRECISION`	convergence tolerance limit for the Fisher-scoring algorithm, Default: `1e-4`
`data`	dataframe containing the variables named in `formula` and `vardir`

Value

The function returns a list with the following objects:

eblup a dataframe with the values of the EBLUP estimators

MSE a dataframe with the estimated mean squared errors of the EBLUPs for the small domains

randomEffect a dataframe with the values of the random effect estimators

Rmatrix a block diagonal matrix composed of sampling errors

fit a list containing the following objects:

- method : type of fitting method, named "REML"
- convergence : a logical value of convergence of Fisher Scoring algorithm
- iterations : number of iterations performed by Fisher-Scoring algorithm
- estcoef : a dataframe with the estimated model coefficient in the first column, their standard error in the second column, the t statistics in the third column, and the p-values of the significance of each coefficient in the last column
- refvar : a dataframe with the estimated random effect variance
- informationFisher : a matrix of information Fisher of Fisher-Scoring algorithm

Examples

```
## Load dataset
data(datasae1)

# Compute EBLUP and MSE of Y1 Y2 and Y3  based on Model 1
# 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", "v2", "v3", "v12", "v13", "v23")
m1 <- eblupMFH1(Fo, vardir, data=datasae1)

## Without parameter 'data'
Fo <- list(f1=datasae1$Y1~datasae1$X1+datasae1$X2,
```

```
          f2=datasae1$Y2~datasae1$X1+datasae1$X2,
          f3=datasae1$Y3~datasae1$X1+datasae1$X2)
vardir <- datasae1[,c("v1", "v2", "v3", "v12", "v13", "v23")]
m1 <- eblupMFH1(Fo, vardir)

m1$eblup   # see the EBLUP estimators
m1$MSE   # see MSE of EBLUP estimators
```

---

eblupMFH2 *EBLUPs based on a Autoregressive Multivariate Fay Herriot (Model 2)*

Description

This function gives the EBLUP and MSE based on a autoregressive multivariate Fay-Herriot model (model 2).

Usage

```
eblupMFH2(formula, vardir, MAXITER = 100, PRECISION = 1e-04, data)
```

Arguments

`formula`	an object of class list of formula, describe the model to be fitted
`vardir`	if data is available, it is vector containing name of sampling variances of direct estimators. if not, it is data frame of sampling variances of direct estimators. The order is : `var1,var2,. ,var(k) ,cov12,. cov1k,cov23,. ,cov(k-1)(k)`
`MAXITER`	maximum number of iterations allowed in the Fisher-scoring algorithm, Default: `100`
`PRECISION`	convergence tolerance limit for the Fisher-scoring algorithm, Default: `1e-4`
`data`	dataframe containing the variables named in `formula` and `vardir`

Value

The function returns a list with the following objects:

eblup a dataframe with the values of the EBLUP estimators

MSE a dataframe with the estimated mean squared errors of the EBLUPs for the small domains

randomEffect a dataframe with the values of the random effect estimators

Rmatrix a block diagonal matrix composed of sampling errors

fit a list containing the following objects:

- method : type of fitting method, named "REML"
- convergence : a logical value of convergence of Fisher Scoring algorithm
- iterations : number of iterations performed by Fisher-Scoring algorithm

- estcoef : a dataframe with the estimated model coefficient in the first column, their standard error in the second column, the t statistics in the third column, and the p-values of the significance of each coefficient in the last column
- refvar : a dataframe with the estimated random effect variance
- rho : a dataframe with the estimated rho of random effect variance and their rho parameter test based on Model 2
- informationFisher : a matrix of information Fisher of Fisher-Scoring algorithm

Examples

```
## Load dataset
data(datasae2)

# Compute EBLUP and MSE of Y1 Y2 and Y3  based on Model 2
# using auxiliary variables X1 and X2 for each dependent variable

## Without parameter 'data'
Fo <- list(f1=Y1~X1+X2,
           f2=Y2~X1+X2,
           f3=Y3~X1+X2)
vardir <- c("v1", "v2", "v3", "v12", "v13", "v23")
m2 <- eblupMFH2(Fo, vardir, data=datasae2)

## Without parameter 'data'
Fo <- list(f1=datasae2$Y1~datasae2$X1+datasae2$X2,
           f2=datasae2$Y2~datasae2$X1+datasae2$X2,
           f3=datasae2$Y3~datasae2$X1+datasae2$X2)
vardir <- datasae2[,c("v1", "v2", "v3", "v12", "v13", "v23")]
m2 <- eblupMFH2(Fo, vardir)

m2$eblup   # see the EBLUP estimators
m2$MSE   # see MSE of EBLUP estimators
```

---

eblupMFH3 — *EBLUPs based on a Heteroscedastic Autoregressive Multivariate Fay Herriot (Model 3)*

Description

This function gives the EBLUP and MSE based on a heteroscedastic autoregressive multivariate Fay-Herriot model (model 3).

Usage

```
eblupMFH3(formula, vardir, MAXITER = 100, PRECISION = 1e-04, data)
```

Arguments

`formula`	an object of class list of formula, describe the model to be fitted
`vardir`	if data is available, it is vector containing name of sampling variances of direct estimators. if not, it is data frame of sampling variances of direct estimators. The order is : `var1,var2,. ,var(k) ,cov12,. cov1k,cov23,. ,cov(k-1)(k)`
`MAXITER`	maximum number of iterations allowed in the Fisher-scoring algorithm, Default: `100`
`PRECISION`	convergence tolerance limit for the Fisher-scoring algorithm, Default: `1e-4`
`data`	dataframe containing the variables named in `formula` and `vardir`

Value

The function returns a list with the following objects:

eblup a dataframe with the values of the EBLUP estimators

MSE a dataframe with the estimated mean squared errors of the EBLUPs for the small domains

randomEffect a dataframe with the values of the random effect estimators

Rmatrix a block diagonal matrix composed of sampling errors

fit a list containing the following objects:

- method : type of fitting method, named "REML"
- convergence : a logical value of convergence of Fisher Scoring algorithm
- iterations : number of iterations performed by Fisher-Scoring algorithm
- estcoef : a dataframe with the estimated model coefficient in the first column, their standard error in the second column, the t statistics in the third column, and the p-values of the significance of each coefficient in the last column
- refvar : a dataframe with the estimated random effect variance
- refvarTest : homogeneity of random effect variance test based on Model 3
- rho : a dataframe with the estimated rho of random effect variance and their rho parameter test based on Model 2
- informationFisher : a matrix of information Fisher of Fisher-Scoring algorithm

Examples

```
## Load dataset
data(datasae3)

# Compute EBLUP and MSE of Y1 Y2 and Y3  based on Model 3
# 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", "v2", "v3", "v12", "v13", "v23")
m3 <- eblupMFH3(Fo, vardir, data=datasae3)
```

```
## Without parameter 'data'
Fo <- list(f1=datasae3$Y1~datasae3$X1+datasae3$X2,
           f2=datasae3$Y2~datasae3$X1+datasae3$X2,
           f3=datasae3$Y3~datasae3$X1+datasae3$X2)
vardir <- datasae3[,c("v1", "v2", "v3", "v12", "v13", "v23")]
m3 <- eblupMFH3(Fo, vardir)

m3$eblup   # see the EBLUP estimators
m3$MSE   # see MSE of EBLUP estimators
```

`eblupUFH` *EBLUPs based on a Univariate Fay Herriot (Model 0)*

Description

This function gives the EBLUP and MSE based on a univariate Fay Herriot model (model 0)

Usage

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

Arguments

`formula`	an object of class list of formula, describe the model to be fitted
`vardir`	if data is available, it is vector containing name of sampling variances of direct estimators. if not, it is data frame of sampling variances of direct estimators. The order is : `var1,var2,. ,var(k) ,cov12,. cov1k,cov23,. ,cov(k-1)(k)`
`samevar`	logical input, true if variance of the data is same, Default: `FALSE`
`MAXITER`	maximum number of iterations allowed in the Fisher-scoring algorithm, Default: `100`
`PRECISION`	convergence tolerance limit for the Fisher-scoring algorithm, Default: `1e-4`
`data`	dataframe containing the variables named in `formula` and `vardir`

Value

The function returns a list with the following objects:

eblup a dataframe with the values of the EBLUP estimators

MSE a dataframe with the estimated mean squared errors of the EBLUPs for the small domains

randomEffect a dataframe with the values of the random effect estimators

Rmatrix a block diagonal matrix composed of sampling errors

fit a list containing the following objects:

- method : type of fitting method, named "REML"
- convergence : a logical value of convergence of Fisher Scoring algorithm
- iterations : number of iterations performed by Fisher-Scoring algorithm
- estcoef : a dataframe with the estimated model coefficient in the first column, their standard error in the second column, the t statistics in the third column, and the p-values of the significance of each coefficient in the last column
- refvar : a dataframe with the estimated random effect variance
- informationFisher : a matrix of information Fisher of Fisher-Scoring algorithm

Examples

```
## Load dataset
data(datasae1)

# Compute EBLUP and MSE of Y1 Y2 and Y3  based on Model 0
# 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", "v2", "v3", "v12", "v13", "v23")
un <- eblupUFH(Fo, vardir, data=datasae1)

## Without parameter 'data'
Fo <- list(f1=datasae1$Y1~datasae1$X1+datasae1$X2,
           f2=datasae1$Y2~datasae1$X1+datasae1$X2,
           f3=datasae1$Y3~datasae1$X1+datasae1$X2)
vardir <- datasae1[,c("v1", "v2", "v3", "v12", "v13", "v23")]
un <- eblupMFH1(Fo, vardir)

un$eblup   # see the EBLUP estimators
un$MSE   # see MSE of EBLUP estimators
```

---

msae *msae : Multivariate Fay Herriot Models for Small Area Estimation*

Description

Implements multivariate Fay-Herriot models for small area estimation. It uses empirical best linear unbiased prediction (EBLUP) estimator. Multivariate models consider the correlation of several target variable and borrow strength from auxiliary variables to improve the effectiveness of a domain sample size. Models which accommodated by this package are univariate model with several target variables (model 0), multivariate model (model 1), autoregressive multivariate model (model 2), and heteroscedastic autoregressive multivariate model (model 3). Functions provide EBLUP estimators and mean squared error (MSE) estimator for each model. These models were developed by Roberto Benavent and Domingo Morales (2015) <doi:10.1016/j.csda.2015.07.013>.

Author(s)

Novia Permatasari, Azka Ubaidillah

Maintainer: Novia Permatasari <16.9335@stis.ac.id>

Functions

`eblupUFH` Gives the EBLUPs and MSE of Univariate SAE (Model 0)

`eblupMFH1` Gives the EBLUPs and MSE of Multivariate SAE (Model 1)

`eblupMFH2` Gives the EBLUPs and MSE of Autoregressive Multivariate SAE (Model 2)

`eblupMFH3` Gives the EBLUPs and MSE of Heteroscedastics Autoregressive Multivariate SAE (Model 3)

Reference

- Benavent, Roberto & Morales, Domingo. (2015). Multivariate Fay-Herriot models for small area estimation. Computational Statistics & Data Analysis. 100. 372-390. DOI: 10.1016/j.csda.2015.07.013.
- Rao, J.N.K & Molina. (2015). Small Area Estimation 2nd Edition. New York: John Wiley and Sons, Inc.
- Ubaidillah, Azka et al. (2019). Multivariate Fay-Herriot models for small area estimation with application to household consumption per capita expenditure in Indonesia. Journal of Applied Statistics. 46:15. 2845-2861. DOI: 10.1080/02664763.2019.1615420.

Index