Package 'saery'

September 10, 2014

Type Package

Title Small Area Estimation for Rao and Yu Model

Version 1.0

Date 2014-09-10

Author Maria Dolores Esteban Lefler, Domingo Morales Gonzalez, Agustin Perez Martin

Maintainer Agustin Perez Martin <agustin.perez@umh.es>

Description A complete set of functions to calculate several EBLUP (Empirical Best Linear Unbiased Predictor) estimators and their mean squared errors. All estimators are based on an area-level linear mixed model introduced by Rao and Yu in 1994 (see documentation). The REML method is used for fitting this model.

License GPL-2

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2014-09-10 23:18:02

R topics documented:

saery-package *Small Area Estimation for Rao and Yu model*

Description

A complete set of functions to calculate several eblups estimators and its mean square errors. All estimators are based in area-level linear mixed model introduced by Rao and Yu in 1994 (see documentation). Saery package are developed to fit the model with REML method.

Details

Package:	saery
Type:	Package
Version:	1.0
Date:	2014-09-10
License:	GPL-2

The main functions of the saery package are `fit.saery` and `eblup.saery`. The function `fit.saery` is used to fit the correct model for three options. `eblup.saery` calculates the eblup and mse for the model.

Author(s)

Maria Dolores Esteban Lefler, Domingo Morales Gonzalez, Agustin Perez Martin

Maintainer: Agustin Perez Martin <agustin.perez@umh.es>

References

Rao, J.N.K., Yu, M., 1994. Small area estimation by combining time series and cross sectional data. Canadian Journal of Statistics 22, 511-528.

Esteban, M.D., Morales, D., Perez, A., Santamaria, L., 2012. Small area estimation of poverty proportions under area-level time models. Computational Statistics and Data Analysis, 56 (10), pp. 2840-2855.

Examples

```
sigma2edi <- datos[,6]
X <- as.matrix(datos[,5])
ydi <- datos[,3]
D <- length(unique(datos[,1]))
md <- rep(length(unique(datos[,2])), D)

output.fit.ar1 <- fit.saery(X, ydi, D, md, sigma2edi, "AR", 0.9)
output.fit.ar1
```

```
#For computational reasons B is too low. We recomend to increase up to 100
eblup.output.ar1 <- eblup.saery(X, ydi, D, md, sigma2edi, "a", plot = TRUE, B = 2)
eblup.output.ar1
```

`datos` *Dataset for saery package*

Description

A simulated dataset created by the authors in order to check the correct operation of the saery package

Usage

```
data(datos)
```

Format

A simulated data frame with 2000 observations on the following 6 variables.

`Area` a numeric vector with the area (domain) of the data

`Period` a numeric vector with the period (subdomain) of the data

`ydi` a numeric vector with the direct estimator of the indicator of interest for area (domain)

`ones` a numeric vector. This is only needed to include the intercept parameter in the model

`xdi` a numeric vector containing the aggregated (population) values of an auxiliary variable

`sigma2edi` a numeric vector with the known variance of the error term

Examples

```
sigma2edi <- datos[,6]
X <- as.matrix(datos[,5])
ydi <- datos[,3]
D <- length(unique(datos[,1]))
md <- rep(length(unique(datos[,2])), D)

output.fit.ar1 <- fit.saery(X, ydi, D, md, sigma2edi, "AR", 0.9)
output.fit.ar1

#For computational reasons B is too low. We recomend to increase up to 100
eblup.output.ar1 <- eblup.saery(X, ydi, D, md, sigma2edi, "a", plot = TRUE, B = 2)
eblup.output.ar1
```

eblup.saery *The function* eblup.saery *calculate the eblup and mse for a model.*

Description

The function eblup.saery calculate the eblup and mse for a model. Is recomended that the model was previusly checked by fit.saery

Usage

```
eblup.saery(X, ydi, D, md, sigma2edi,
      model = c("INDEP", "AR1", "MA1"),
      plot = FALSE, type = "I", B = 100)
eblup.saery.AR1(X, ydi, D, md, sigma2edi, plot, type, B)
eblup.saery.MA1(X, ydi, D, md, sigma2edi, plot, type, B)
eblup.saery.indep(X, ydi, D, md, sigma2edi, plot)
```

Arguments

X	a numeric vector or data frame containing the aggregated (population) values of p auxiliary variables. A ones columns must be agregated to calculate the intercept parameter
ydi	a numeric vector with the direct estimator of the indicator of interest for area (domain)
D	a numeric vector with the number of areas (domain) of the data
md	a numeric vector with the number of periods (subdomains) for each area of the data
sigma2edi	a numeric vector with the known variance of the error term
model	Three diferents types of model must be fit. For an indepent model INDEP or an abbreviated version of it must be typed. An AR1 model can be fit using AR1 or an abbreviated version. For MA1 model, use MA1 or an abbreviated version. This function and fit.saery use the REML method to fit the model.
plot	logical specifying if a set of plot be returned. FALSE by default.
type	three types of mse can be calculated for AR1 and MA1 models. One of the with an explicit and direct function, and the orther two by a bootstrap procedure that takes more time to the system to calculate them. For the indep model, only the an explicit function are developed type = "I"
B	the number of bootstrap samples to be generated and fitted for types "I" and "II"

Value

A data frame with the eblups and its mse be returned. A set of plots be displayed if plot = TRUE

Author(s)

Maria Dolores Esteban Lefler, Domingo Morales Gonzalez, Agustin Perez Martin

References

Rao, J.N.K., Yu, M., 1994. Small area estimation by combining time series and cross sectional data. Canadian Journal of Statistics 22, 511-528.

Esteban, M.D., Morales, D., Perez, A., Santamaria, L., 2012. Small area estimation of poverty proportions under area-level time models. Computational Statistics and Data Analysis, 56 (10), pp. 2840-2855.

See Also

fit.saery

Examples

```
sigma2edi <- datos[,6]
X <- as.matrix(datos[,5])
ydi <- datos[,3]
D <- length(unique(datos[,1]))
md <- rep(length(unique(datos[,2])), D)

#For computational reasons B is too low. We recomend to increase up to 100
eblup.output.ar1 <- eblup.saery(X, ydi, D, md, sigma2edi, model = "a", plot = TRUE, B = 2)
eblup.output.ar1

#For computational reasons B is too low. We recomend to increase up to 100
eblup.output.ma1 <- eblup.saery(X, ydi, D, md, sigma2edi,
                    model = "ma", plot = FALSE, type = "II", B = 2)
eblup.output.ma1

eblup.output.indep <- eblup.saery(X, ydi, D, md, sigma2edi,
                      model = "i", plot = TRUE)
eblup.output.indep
```

`fit.saery` *The function* `fit.saery` *is used to fit the correct model for three options.*

Description

The function `fit.saery` fits the model for three options. This function and `eblup.saery` use the REML method to fit the model.

Usage

```
fit.saery(X, ydi, D, md, sigma2edi,
      model = c("INDEP", "AR1", "MA1"),
      conf.level = 0.95)
fit.saery.AR1(X, ydi, D, md, sigma2edi, conf.level)
fit.saery.MA1(X, ydi, D, md, sigma2edi, conf.level)
fit.saery.indep(X, ydi, D, md, sigma2edi, conf.level)
```

Arguments

X	a numeric vector or data frame containing the aggregated (population) values of p auxiliary variables. A ones columns must be agregated to calculate the intercept parameter
ydi	a numeric vector with the direct estimator of the indicator of interest for area (domain)
D	a numeric vector with the number of areas (domain) of the data
md	a numeric vector with the number of periods (subdomains) for each area of the data
sigma2edi	a numeric vector with the known variance of the error term
model	Three diferents types of model must be fit. For an indepent model INDEP or an abbreviated version of it must be typed. An AR1 model can be fit using AR1 or an abbreviated version. For MA1 model, use MA1 or an abbreviated version. This function and fit.saery use the REML method to fit the model.
conf.level	a value under 1 for the confidence level for the confidence intervals returned by the function

Value

A list with the fitted parameters of the model are returned. Caonfidence intervals, p-values, the Fisher Scoring matrix and the number of iterations of the model are also returned.

Author(s)

Maria Dolores Esteban Lefler, Domingo Morales Gonzalez, Agustin Perez Martin

References

Rao, J.N.K., Yu, M., 1994. Small area estimation by combining time series and cross sectional data. Canadian Journal of Statistics 22, 511-528.

Esteban, M.D., Morales, D., Perez, A., Santamaria, L., 2012. Small area estimation of poverty proportions under area-level time models. Computational Statistics and Data Analysis, 56 (10), pp. 2840-2855.

See Also

eblup.saery, ~~~

Examples

```
sigma2edi <- datos[,6]
X <- as.matrix(datos[,5])
ydi <- datos[,3]
D <- length(unique(datos[,1]))
md <- rep(length(unique(datos[,2])), D)

output.fit.ar1 <- fit.saery(X, ydi, D, md, sigma2edi, "AR", 0.9)
output.fit.ar1

output.fit.ma1 <- fit.saery(X, ydi, D, md, sigma2edi, "MA", 0.9)
output.fit.ma1

output.fit.indep <- fit.saery(X, ydi, D, md, sigma2edi, "indep", 0.9)
output.fit.indep
```

Index