Package ‘BayesSampling’

April 24, 2020

Type Package
Title Bayes Linear Estimators for Finite Population
Version 1.0.0
Date 2020-04-13
Maintainer Pedro Soares Figueiredo <pedrosfig@hotmail.com>
Description Allows the user to apply the Bayes Linear approach to finite population with the Simple Random Sampling - BLE_SRS() - and the Stratified Simple Random Sampling design - BLE_SSRS() - (both without replacement) and to the Ratio estimator (using auxiliary information) - BLE_Ratio().
The Bayes linear estimation approach is applied to a general linear regression model for finite population prediction in BLE_Reg()
and it is also possible to achieve the design based estimators using vague prior distributions.
URL https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X201400111886,
https://github.com/pedrosfig/BayesSampling
License GPL-3
Encoding UTF-8
LazyData true
RoxygenNote 7.0.2
Depends R (>= 3.5)
Imports MASS, Matrix, stats
Suggests knitr, rmarkdown, TeachingSampling
VignetteBuilder knitr
Language en-US
NeedsCompilation no
Author Pedro Soares Figueiredo [aut, cre]
Kelly C. M. Gonçalves [aut, ths]
(<https://orcid.org/0000-0003-2279-2881>),
(<https://orcid.org/0000-0002-4524-547X>)
Description

This data set corresponds to some socioeconomic variables from 150266 people of a city in a particular year.

Usage

```r
data(BigCity)
```

Format

A data.frame with 150266 rows and 12 variables:

- **HHID** The identifier of the household. It corresponds to an alphanumeric sequence (four letters and five digits).
- **PersonID** The identifier of the person within the household. NOTE it is not a unique identifier of a person for the whole population. It corresponds to an alphanumeric sequence (five letters and two digits).
- **Stratum** Households are located in geographic strata. There are 119 strata across the city.
- **PSU** Households are clustered in cartographic segments defined as primary sampling units (PSU). There are 1664 PSU and they are nested within strata.
- **Zone** Segments clustered within strata can be located within urban or rural areas along the city.
BLE_Ratio

Sex  Sex of the person.
Income  Per capita monthly income.
Expenditure  Per capita monthly expenditure.
Employment  A person’s employment status.
Poverty  This variable indicates whether the person is poor or not. It depends on income.

Source

https://CRAN.R-project.org/package=TeachingSampling

References

Package ‘TeachingSampling’: see BigCity

BLE_Ratio  Ratio BLE

Description

Creates the Bayes Linear Estimator for the Ratio "estimator"

Usage

BLE_Ratio(ys, xs, x_nots, m = NULL, v = NULL, sigma = NULL)

Arguments

ys  vector of sample observations.
xs  vector with values for the auxiliary variable of the elements in the sample.
x_nots  vector with values for the auxiliary variable of the elements not in the sample.
m  prior mean for the ratio between Y and X. If NULL, mean(ys)/mean(xs) will be used (non-informative prior).
v  prior variance of the ratio between Y and X (bigger than sigma^2). If NULL, it will tend to infinity (non-informative prior).
sigma  prior estimate of variability (standard deviation) of the ratio within the population. If NULL, sample variance of the ratio will be used.

Value

A list containing the following components:

• est.beta - BLE of Beta
• Vest.beta - Variance associated with the above
• est.mean - BLE for each individual not in the sample
• Vest.mean - Covariance matrix associated with the above
• est.tot - BLE for the total
• Vest.tot - Variance associated with the above
References


Examples

ys <- c(10,8,6)
xs <- c(5,4,3)
x_nots <- c(1,20)
m <- 2
v <- 10
sigma <- 2

Estimator <- BLE_Ratio(ys,xs,x_nots,m,v,sigma)
Estimator

Description

Calculates the Bayes Linear Estimator for Regression models (general case)

Usage

BLE_Reg(ys, xs, a, R, Vs, x_nots, V_nots)

Arguments

ys response variable of the sample
xs explicative variable of the sample
a vector of means from Beta
R covariance matrix of Beta
Vs covariance of sample errors
x_nots values of X for the individuals not in the sample
V_nots covariance matrix of the individuals not in the sample
**Value**

A list containing the following components:

- est.beta - BLE of Beta
- Vest.beta - Variance associated with the above
- est.mean - BLE of each individual not in the sample
- Vest.mean - Covariance matrix associated with the above
- est.tot - BLE for the total
- Vest.tot - Variance associated with the above

**Source**


**References**


**Examples**

```r
xs <- matrix(c(1,1,1,1,2,3,5,0),nrow=4,ncol=2)
y <- c(12,17,28,2)
x_nots <- matrix(c(1,1,1,0,1,4),nrow=3,ncol=2)
a <- c(1.5,6)
R <- matrix(c(10,2,2,10),nrow=2,ncol=2)
V < - diag(c(1,1,1))
V_nots < - diag(c(1,1,1))

Estimator <- BLE_Reg(ys,xs,a,R,V,x_nots,V_nots)
Estimator
```

**BLE_SRS**

**Simple Random Sample BLE**

**Description**

Creates the Bayes Linear Estimator for the Simple Random Sampling design (without replacement)

**Usage**

```r
BLE_SRS(ys, N, m = NULL, v = NULL, sigma = NULL)
```
Arguments

- **ys**: vector of sample observations.
- **N**: total size of the population.
- **m**: prior mean. If NULL, sample mean will be used (non-informative prior).
- **v**: prior variance of an element from the population (bigger than sigma^2). If NULL, it will tend to infinity (non-informative prior).
- **sigma**: prior estimate of variability (standard deviation) within the population. If NULL, sample variance will be used.

Value

A list containing the following components:

- **est.beta**: BLE of Beta (BLE for every individual)
- **Vest.beta**: Variance associated with the above
- **est.mean**: BLE for each individual not in the sample
- **Vest.mean**: Covariance matrix associated with the above
- **est.tot**: BLE for the total
- **Vest.tot**: Variance associated with the above

Source


References


Examples

```r
ys <- c(5,6,8)
m <- 6
v <- 5
sigma <- 1
N <- 5

Estimator <- BLE_SRS(ys,N,m,v,sigma)
Estimator
```
Description

Creates the Bayes Linear Estimator for the Stratified Simple Random Sampling design (without replacement)

Usage

BLE_SSRS(ys, h, N, m = NULL, v = NULL, sigma = NULL)

Arguments

ys vector of sample observations.

h vector with number of observations in each strata.

N vector with the total size of each strata.

m vector with the prior mean of each strata. If NULL, sample mean for each strata will be used (non-informative prior).

v vector with the prior variance of an element from each strata (bigger than sigma^2 for each strata). If NULL, it will tend to infinity (non-informative prior).

sigma vector with the prior estimate of variability (standard deviation) within each strata of the population. If NULL, sample variance of each strata will be used.

Value

A list containing the following components:

• est.beta - BLE of Beta (BLE for the individuals in each strata)
• Vest.beta - Variance associated with the above
• est.mean - BLE for each individual not in the sample
• Vest.mean - Covariance matrix associated with the above
• est.tot - BLE for the total
• Vest.tot - Variance associated with the above

Source


References

Examples

ys <- c(2,-1,1.5, 6,10, 8,8)
h <- c(3,2,2)
m <- c(0,9,8)
v <- c(3,8,1)
sigma <- c(1,2,0.5)
N <- c(5,5,3)

Estimator <- BLE_SSRS(ys,h,N,m,v,sigma)

C calculates the C factor

Description

calculates the C factor

Usage

C(ys, xs, R, Vs)

Arguments

ys response variable of the sample
xs explicative variable of the sample
R covariance matrix of Beta
Vs covariance of sample errors

create1 creates vector of 1’s to be used in the estimators

Description

creates vector of 1’s to be used in the estimators

Usage

create1(y)

Arguments

y sample matrix

Value

vector of 1’s with size equal to the number of observations in the sample
\textit{E\textsubscript{\textit{beta}}} \quad \textit{calculates the BLE for Beta}

**Description**

calculates the BLE for Beta

**Usage**

\texttt{E\_beta(ys, xs, a, R, Vs)}

**Arguments**

- \texttt{ys} \quad \textit{response variable of the sample}
- \texttt{xs} \quad \textit{explicative variable of the sample}
- \texttt{a} \quad \textit{vector of means from Beta}
- \texttt{R} \quad \textit{covariance matrix of Beta}
- \texttt{Vs} \quad \textit{covariance of sample errors}

\textit{E\textsubscript{\textit{theta\_Reg}}} \quad \textit{calculates the BLE for the individuals not in the sample}

**Description**

calculates the BLE for the individuals not in the sample

**Usage**

\texttt{E\_theta\_Reg(ys, xs, a, R, Vs, x\_not)}

**Arguments**

- \texttt{ys} \quad \textit{response variable of the sample}
- \texttt{xs} \quad \textit{explicative variable of the sample}
- \texttt{a} \quad \textit{vector of means from Beta}
- \texttt{R} \quad \textit{covariance matrix of Beta}
- \texttt{Vs} \quad \textit{covariance of sample errors}
- \texttt{x\_not} \quad \textit{values of X for the individuals not in the sample}
VT_Reg

**Description**
calculates BLE for the total $T$

**Usage**
$T\_Reg(ys, xs, a, R, Vs, x\_nots)$

**Arguments**
- $ys$: response variable of the sample
- $xs$: explicative variable of the sample
- $a$: vector of means from Beta
- $R$: covariance matrix of Beta
- $Vs$: covariance of sample errors
- $x\_nots$: values of $X$ for the individuals not in the sample

VT_Reg

**Description**
calculates risk matrix associated with the BLE for for the total $T$

**Usage**
$VT\_Reg(ys, xs, a, R, Vs, x\_nots, V\_nots)$

**Arguments**
- $ys$: response variable of the sample
- $xs$: explicative variable of the sample
- $a$: vector of means from Beta
- $R$: covariance matrix of Beta
- $Vs$: covariance of sample errors
- $x\_nots$: values of $X$ for the individuals not in the sample
- $V\_nots$: covariance matrix of the individuals not in the sample
**V_beta**

*calculates the risk matrix associated with the BLE for Beta*

**Description**

calculates the risk matrix associated with the BLE for Beta

**Usage**

\[
V_{\text{beta}}(y_s, x_s, R, V_s)
\]

**Arguments**

- **y_s**: response variable of the sample
- **x_s**: explicative variable of the sample
- **R**: covariance matrix of Beta
- **V_s**: covariance of sample errors

---

**V_theta_Reg**

*calculates the risk matrix associated with the BLE for the individuals not in the sample*

**Description**

calculates the risk matrix associated with the BLE for the individuals not in the sample

**Usage**

\[
V_{\text{theta}_\text{Reg}}(y_s, x_s, R, V_s, x_{\text{nots}}, V_{\text{nots}})
\]

**Arguments**

- **y_s**: response variable of the sample
- **x_s**: explicative variable of the sample
- **R**: covariance matrix of Beta
- **V_s**: covariance of sample errors
- **x_{\text{nots}}**: values of X for the individuals not in the sample
- **V_{\text{nots}}**: covariance matrix of the individuals not in the sample
Index

*Topic **datasets**
  BigCity, 2

BigCity, 2, 3
BLE_Ratio, 3
BLE_Reg, 4
BLE_SRS, 5
BLE_SSRS, 7

C, 8
create1, 8

E_beta, 9
E_theta_Reg, 9

T_Reg, 10

V_beta, 11
V_theta_Reg, 11
VT_Reg, 10