Package ‘saeHB.panel.beta’

June 20, 2023

Type Package

Title Small Area Estimation using HB for Rao Yu Model under Beta Distribution

Version 0.1.2

Maintainer Dian Rahmawati Salis <dianrahmawatisalis03@gmail.com>

Description Several functions are provided for small area estimation at the area level using the hierarchical bayesian (HB) method with panel data under beta distribution for variable interest. This package also provides a dataset produced by data generation. The 'rjags' package is employed to obtain parameter estimates. Model-based estimators involve the HB estimators, which include the mean and the variation of the mean. For the reference, see Rao and Molina (2015, ISBN: 978-1-118-73578-7).

License GPL-3

URL https://github.com/DianRahmawatiSalis/saeHB.panel.beta

BugReports https://github.com/DianRahmawatiSalis/saeHB.panel.beta/issues

Depends R(>= 2.10)

Imports coda, dplyr, graphics, grDevices, rjags, stats, stringr

Suggests knitr, R.rsp, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr, R.rsp

Config/testthat/edition 3

Encoding UTF-8

LazyData true

NeedsCompilation no

RoxygenNote 7.2.3

SystemRequirements JAGS (http://mcmc-jags.sourceforge.net)

Author Dian Rahmawati Salis [aut, cre],
Azka Ubaidillah [aut]

Repository CRAN

Date/Publication 2023-06-20 05:00:02 UTC
Description

Dataset under Beta Distribution to simulate Small Area Estimation using Hierarchical Bayesian Method for Rao Yu Model This data is generated by these following steps:

1. Generate random effect area $v$, random effect for area $i$ at time point $j$ $u$, epsilon $\epsilon$, variance of $y_{di}$ $\text{vardir}_{ij}$, sampling error $e$, auxiliary $x_{di1}$ and $x_{di2}$
   
   - Set coefficient $\beta_0 = \beta_1 = \beta_2 = 2$ and $\rho = -0.5$
   - Generate random effect area $v_{(i)} \sim N(0,1)$
   - Generate auxiliary variable $x_{di1}_{(ij)} \sim U(0,1)$
   - Generate auxiliary variable $x_{di2}_{(ij)} \sim U(0,1)$
   - Generate epsilon $\epsilon_{ij} \sim N(0,1)$
   - Generate sampling error $e_{ij} \sim N(0,\text{vardir}_{ij})$
   - Generate $\phi_{ij} \sim \text{Gamma}(1,0.5)$
   - Calculate random effect for area $i$ at time point $j$ $u_{ij} = \rho * u_{ij-1} + \epsilon_{ij}$
   - Calculate $\mu_{ij} = \frac{(\exp \beta_0 + \beta_1 x_{di1_{ij}} + \beta_2 x_{di2_{ij}} + v_{i} + \epsilon_{ij})}{(1 + \exp \beta_0 + \beta_1 x_{di1_{ij}} + \beta_2 x_{di2_{ij}} + v_{i} + \epsilon_{ij})}$
   - Calculate $A_{ij} = \mu_{ij} * \phi_{ij}$
   - Calculate $B_{ij} = (1 - \mu_{ij}) * \phi_{ij}$
   - Generate $y_{ij} \sim \text{Beta}(A_{ij}, B_{ij})$
   - Calculate variance of $y_{di}$ with $\text{vardir}_{ij} = \frac{(A_{ij})(B_{ij})}{(A_{ij} + B_{ij})^2(A_{ij} + B_{ij} + 1)}$
   - Set area=20 and period=5

2. Auxiliary variables $x_{di1}$, $x_{di2}$, direct estimation $y$, area, period, and $\text{vardir}$ are combined in a dataframe called dataAr1

Usage

dataBetaAr1
Format
A data frame with 100 rows and 6 variables:

- **ydi**: Direct Estimation of y
- **area**: Area (domain) of the data
- **period**: Period (subdomain) of the data
- **vardir**: Sampling Variance of y
- **xdi1**: Auxiliary variable of xdi1
- **xdi2**: Auxiliary variable of xdi2

---

**Description**

1. A dataset under Beta Distribution to simulate Small Area Estimation using Hierarchical Bayesian method for Rao-Yu Model with Non-sampled Area
2. This data contains NA values that indicates no sampled in at least one area.

**Usage**

dataBetaAr1Ns

**Format**
A data frame with 100 row and 6 column:

- **ydi**: Direct Estimation of y
- **area**: Area (domain) of the data
- **period**: Period (subdomain) of the data
- **vardir**: Sampling Variance of y
- **xdi1**: Auxiliary variable of xdi1
- **xdi2**: Auxiliary variable of xdi2
Dataset under Beta Distribution to simulate Small Area Estimation using Hierarchical Bayesian Method for Rao-Yu Model with \( \rho = 0 \) This data is generated by these following steps:

1. Generate random effect area \( v \), random effect for area \( i \) at time point \( j \), \( u \), \( \epsilon \), variance of \( y_{ij} \) \( \text{vardir}_{ij} \), sampling error \( e \), auxiliary \( x_{di1} \) and \( x_{di2} \)
   - Set coefficient \( \beta_0 = \beta_1 = \beta_2 = 2 \)
   - Generate random effect area \( v_{i} \sim \text{N}(0,1) \)
   - Generate auxiliary variable \( x_{di1}_{ij} \sim \text{U}(0,1) \)
   - Generate auxiliary variable \( x_{di2}_{ij} \sim \text{U}(0,1) \)
   - Generate epsilon \( \epsilon_{ij} \sim \text{N}(0,1) \)
   - Generate \( \phi_{ij} \sim \text{Gamma}(1,0.5) \)
   - Calculate \( \mu_{ij} = \exp(\beta_0 + \beta_1 x_{di1_{ij}} + \beta_2 x_{di2_{ij}} + v_i + \epsilon_{ij}) \)
   - Calculate \( A_{ij} = \mu_{ij} \times \phi_{ij} \)
   - Calculate \( B_{ij} = (1 - \mu_{ij}) \times \phi_{ij} \)
   - Generate \( y_{ij} \sim \text{Beta}(A_{ij},B_{ij}) \)
   - Calculate variance of \( y_{ij} \) with \( \text{vardir}_{ij} = \frac{(A_{ij} + B_{ij})(A_{ij})}{(A_{ij} + B_{ij} + 1)} \)
   - Set area=20 and period=5

2. Auxiliary variables \( x_{di1}, x_{di2} \), direct estimation \( y \), area, period, and \( \text{vardir} \) are combined in a dataframe called \( \text{dataPanel} \)

**Usage**

\( \text{dataPanelbeta} \)

**Format**

A data frame with 100 rows and 6 variables:

- \( y_{di} \) Direct Estimation of \( y \)
- \( \text{area} \) Area (domain) of the data
- \( \text{period} \) Period (subdomain) of the data
- \( \text{vardir} \) Sampling Variance of \( y \)
- \( x_{di1} \) Auxiliary variable of \( x_{di1} \)
- \( x_{di2} \) Auxiliary variable of \( x_{di2} \)
dataPanelbetaNs

Sample Data under Beta Distribution for Small Area Estimation using Hierarchical Bayesian Method for Rao Yu Model when rho = 0 with Non Sampled Area

Description

1. A dataset under Beta Distribution to simulate Small Area Estimation using Hierarchical Bayesian method for Rao-Yu Model with Non-sampled area
2. This data contains NA values that indicates no sampled in at least one area.

Usage
dataPanelbetaNs

Format

A data frame with 100 row and 6 column:

- **ydi**: Direct Estimation of y
- **area**: Area (domain) of the data
- **period**: Period (subdomain) of the data
- **vardir**: Sampling Variance of y
- **xdi1**: Auxiliary variable of xdi1
- **xdi2**: Auxiliary variable of xdi2

Panel.beta

Small Area Estimation using Hierarchical Bayesian for Rao-Yu Model under Beta Distribution with rho=0

Description

This function is implemented to variable of interest ydi

Usage

Panel.beta(
  formula,
  area,
  period,
  iter.update = 3,
  iter.mcmc = 2000,
  thin = 1,
  burn.in = 1000,
\begin{verbatim}
Panel.beta

tau.e = 1,
tau.v = 1,
data
)

Arguments

formula  Formula that describe the fitted model
area     Number of areas (domain) of the data
period   Number of periods (subdomains) for each area of the data
iter.update Number of updates with default 3
iter.mcmc Number of total iterations per chain with default 2000
thin     Thinning rate, must be a positive integer with default 1
burn.in  Number of iterations to discard at the beginning with default 1000
tau.e    Variance of area-by-time effect of variable interest with default 1
tau.v    Variance of random area effect of variable interest with default 1
data     The data frame

Value

This function returns a list of the following objects:

Est        A vector with the values of Small Area mean Estimates using Hierarchical bayesian method
refVar     Estimated random effect variances
coeff      A dataframe with the estimated model coefficient
plot       Trace, Density, Autocorrelation Function Plot of MCMC samples
convergence.test    Convergence diagnostic for Markov chains based on Geweke test

Examples

### For data without any non-sampled area
data(dataPanelbeta)  # Load dataset
dataPanelbeta = dataPanelbeta[1:25,]  # for the example only use part of the dataset
formula = ydi ~ xdi1 + xdi2
area = max(dataPanelbeta[, "area"])
period = max(dataPanelbeta[, "period"])

result <- Panel.beta(formula, area, period, data = dataPanelbeta)

result$Est
result$refVar
result$coeff
result$plot

### For data with non-sampled area use dataPanelbetaNs
\end{verbatim}
Description

This function is implemented to variable of interest \( y_{di} \)

Usage

```r
RaoYuAr1.beta(
  formula,  
  area,      
  period,   
  iter.update = 3,  
  iter.mcmc = 2000,  
  thin = 1,  
  burn.in = 1000,  
  tau.e = 1,  
  tau.v = 1,  
  data)
```

Arguments

- `formula`: Formula that describe the fitted model
- `area`: Number of areas (domain) of the data
- `period`: Number of periods (subdomains) for each area of the data
- `iter.update`: Number of updates with default 3
- `iter.mcmc`: Number of total iterations per chain with default 2000
- `thin`: Thinning rate, must be a positive integer with default 1
- `burn.in`: Number of iterations to discard at the beginning with default 1000
- `tau.e`: Variance of area-by-time effect of variable interest with default 1
- `tau.v`: Variance of random area effect of variable interest with default 1
- `data`: The data frame

Value

This function returns a list of the following objects:

- `Est`: A vector with the values of Small Area mean Estimates using Hierarchical bayesian method
- `refVar`: Estimated random effect variances
- `coefficient`: A dataframe with the estimated model coefficient
alpha Parameter dispersion of Generalized Poisson distribution
plot Trace, Density, Autocorrelation Function Plot of MCMC samples
convergence.test Convergence diagnostic for Markov chains based on Geweke test

Examples

```r
# For data without any non-sampled area
data(dataBetaAr1) # Load dataset
dataBetaAr1 = dataBetaAr1[1:25,] # for the example only use part of the dataset
formula = ydi ~ xdi1 + xdi2
area = max(dataBetaAr1[, "area"])
period = max(dataBetaAr1[, "period"])
result <- RaoYuAr1.beta(formula, area, period, data = dataBetaAr1)
result$Est
result$refVar
result$coefficient
result$plot
```

## For data with non-sampled area use dataBetaAr1Ns
Index

* datasets
  dataBetaAr1, 2
  dataBetaAr1Ns, 3
  dataPanelbeta, 4
  dataPanelbetaNs, 5

dataBetaAr1, 2
dataBetaAr1Ns, 3
dataPanelbeta, 4
dataPanelbetaNs, 5

Panel.beta, 5
RaoYuAr1.beta, 7