

# Package ‘saeHB.TF.beta’

May 9, 2026

**Title** SAE using HB Twofold Subarea Model under Beta Distribution

**Version** 0.2.0

**Maintainer** Nasya Zahira Putri <nasyazp28@gmail.com>

**Description** Estimates area and subarea level proportions using the Small Area Estimation (SAE) Twofold Subarea Model with a hierarchical Bayesian (HB) approach under Beta distribution. A number of simulated datasets generated for illustration purposes are also included. The 'rstan' package is employed to estimate parameters via the Hamiltonian Monte Carlo and No U-Turn Sampler algorithm. The model-based estimators include the HB mean, the variation of the mean, and quantiles. For references, see Rao and Molina (2015) <[doi:10.1002/9781118735855](https://doi.org/10.1002/9781118735855)>, Torabi and Rao (2014) <[doi:10.1016/j.jmva.2014.02.001](https://doi.org/10.1016/j.jmva.2014.02.001)>, hadjer et al.(2007) <<http://www.asasrms.org/Proceedings/y2007/Files/JSM2007-000559.pdf>>, Erciulescu et al.(2019) <[doi:10.1111/rssa.12390](https://doi.org/10.1111/rssa.12390)>, and Yudasena (2024).

**License** GPL (>= 3)

**Encoding** UTF-8

**RoxygenNote** 7.3.2

**Biarch** true

**Depends** R (>= 3.5)

**Imports** methods, Rcpp (>= 0.12.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), rstantools (>= 2.4.0), bayesplot, stringr

**LinkingTo** BH (>= 1.66.0), Rcpp (>= 0.12.0), RcppEigen (>= 0.3.3.3.0), RcppParallel (>= 5.0.1), rstan (>= 2.18.1), StanHeaders (>= 2.18.0)

**SystemRequirements** GNU make

**URL** <https://github.com/Nasyazahira/saeHB.TF.beta>

**BugReports** <https://github.com/Nasyazahira/saeHB.TF.beta/issues>

**LazyData** true

**Suggests** knitr, rmarkdown, testthat (>= 3.0.0), ggplot2

**VignetteBuilder** knitr

**Config/testthat/edition** 3

**NeedsCompilation** yes

**Author** Nasya Zahira Putri [aut, cre],  
Cucu Sumarni [aut],  
Rizal Fauziarochman Yudasena [aut]

**Repository** CRAN

**Date/Publication** 2025-09-03 06:50:02 UTC

## Contents

saeHB.TF.beta-package . . . . .	2
betaTF . . . . .	2
dataBeta . . . . .	4
dataBetaNS . . . . .	5
explore . . . . .	6
<b>Index</b>	<b>7</b>

---

saeHB.TF.beta-package *The 'saeHB.TF.beta' package.*

---

## Description

Small Area Estimation using Hierarchical Bayes Twofold Subarea Level Model under Beta Distribution

## References

Stan Development Team (NA). RStan: the R interface to Stan. R package version 2.36.0.9000.  
<https://mc-stan.org>

---

betaTF *Small Area Estimation using Hierarchical Bayes Twofold Subarea Level Model under Beta Distribution*

---

## Description

Function betaTF used for estimation of subarea and area means simultaneously under Twofold Subarea Level Small Area Estimation Model Using Hierarchical Bayesian Method with Beta distribution The range of data must be  $0 < y < 1$ .

**Usage**

```
betaTF(
  formula,
  area,
  weight,
  iter.update = 3,
  iter.mcmc = 1000,
  coef = NULL,
  var.coef = NULL,
  thin = 1,
  burn.in = floor(iter.mcmc/2),
  sigma2.u = 1,
  sigma2.v = 1,
  data
)
```

**Arguments**

formula	Formula that describe the fitted model
area	Index that describes the code relating to area in each subarea. This should be defined for aggregation to get area estimator. Index start from 1 until m
weight	Vector contain proportion units or proportion of population on each subarea. $w_{ij}$
iter.update	Number of updates perform ( default = 3)
iter.mcmc	Number of total iterations per chain (default = 1000)
coef	Vector contains prior initial value of Coefficient of Regression Model for fixed effect with default vector of 0 with the length of the number of regression coefficients
var.coef	Vector contains prior initial value of variance of Coefficient of Regression Model for fixed effect with default vector of 1 with the length of the number of regression coefficients
thin	Thinning rate, must be a positive integer
burn.in	Number of iterations to discard at the beginning
sigma2.u	Number of prior initial value of variance of subarea random effect
sigma2.v	Number of prior initial value of variance of area random effect
data	The data frame

**Value**

This function returns a list with following objects:

Est_sub	A dataframe contains the values, standard deviation, and quantile of Subarea mean Estimates using Twofold Subarea level model under Hierarchical Bayes method
Est_area	A dataframe contains the values, standard deviation, and quantile of Area mean Estimates using Twofold Subarea level model under Hierarchical Bayes method

area_ranreff	A dataframe contains area random effect
sub_ranreff	A dataframe contains subarea random effect
refVar	A dataframe that contains estimated subarea and area random effect variance ( $\sigma_u^2$ and $\sigma_v^2$ )
coefficient	A dataframe that contains the estimated model coefficient $\beta$
plot	Trace, Density, Autocorrelation Function Plot of coefficient

### Examples

```
fit <- betaTF(y~X1+X2,area="codearea",weight="w",data=dataBeta, iter.mcmc = 500)
```

---

dataBeta	<i>Simulated dataset Under Two Fold Subarea level model with Beta distribution.</i>
----------	---

---

### Description

A dataset to simulate Small Area Estimation using Hierarchical Bayesian method under Two Fold Subarea level model with Beta distribution on variable interest.

This data is generated by these following steps:

1. Generate auxiliary variable  $X_{ij1}$ ,  $X_{ij2}$ , sampling error  $e_{ij}$ , subarea random effect  $u_{ij}$ , area random effect  $v_i$ , and weight or proportions of unit  $w_{ij}$ 
  - Generate auxiliary variable on subarea level  $X_{ij1} \sim U(0, 1)$
  - Generate auxiliary variable on subarea level  $X_{ij2} \sim N(0, 1)$
  - Setting coefficient  $\beta_0 = \beta_1 = \beta_2 = 0.5$
  - Generate area random effect  $v_i \sim N(0, 1)$
  - Generate subarea random effect  $u_{ij} \sim N(0, 1)$
  - Calculate target parameter  $\mu_{ij} = \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + v_i + u_{ij}$
  - Generate constant for Beta parameter  $\pi_{ij} \sim \text{Gamma}(1, 0.5)$
  - Calculate Beta parameter  $A = \mu_{ij} \pi_{ij}$  and  $A = (1 - \mu_{ij}) \pi_{ij}$
  - Generate direct estimator  $y_{ij} \sim \text{Beta}(A, B)$
  - Generate weight on each subarea  $w_{ij} \sim U(0.2, 0.7)$
2. Direct estimation ( $y_{ij}$ ), Auxiliary variables  $X_{ij1}$ ,  $X_{ij2}$ , vardir, codearea, and weight  $w_{ij}$  are combined in a dataframe called dataBeta

### Usage

dataBeta

**Format**

A data frame with 90 rows and 6 columns:

**y** Direct estimation of subarea mean  $y_{ij}$

**X1** Auxiliary variable of  $X_{ij1}$

**X2** Auxiliary variable of  $X_{ij2}$

**codearea** Index that describes the code relating to area for each subarea

**w** Unit proportion on each subarea or weight  $w_{ij}$

**vardir** Sampling variance of direct estimator  $y_{ij}$

---

dataBetaNS

*Simulated dataset Under Two Fold Subarea level model with Beta distribution and Non-Sampled subarea.*

---

**Description**

1. A dataset to simulate Small Area Estimation using Hierarchical Bayesian method under Two Fold Subarea level model with Beta distribution and Non-sampled subarea
2. This data contains NA values that indicates no sampled at one or more Subareas. It uses the [dataBeta](#) with the direct estimates and the related variances in 5 subareas are missing.

**Usage**

dataBetaNS

**Format**

A data frame with 90 rows and 6 columns:

**y** Direct estimation of subarea mean  $y_{ij}$

**X1** Auxiliary variable of  $X_{ij1}$

**X2** Auxiliary variable of  $X_{ij2}$

**codearea** Index that describes the code relating to area for each subarea

**w** Unit proportion on each subarea or weight  $w_{ij}$

**vardir** Sampling variance of direct estimator  $y_{ij}$

---

explore

*Exploration of the Data Used for Modeling*

---

### Description

Function `explore` provides an initial exploration of a dataset. It calculate summary statistics for all variables in the provided formula or dataset, visualizes the distribution of the response variable as a histogram density, and boxplot for Coefficient of Variation (CV) / Relative Standard Error (RSE).

### Usage

```
explore(formula, CV = NULL, data, normality = FALSE)
```

### Arguments

<code>formula</code>	Optional formula to specify a response variable (e.g., $y \sim x1 + x2$ ).
<code>CV</code>	Coefficient of Variation (CV) or Relative Standard Error (RSE) of the response variable
<code>data</code>	The dataframe to be explored
<code>normality</code>	Logical; if TRUE, the function will additionally check the normality of the response variable and display the result. Defaults to FALSE.

### Value

Prints a data frame of summary statistics for the selected variables, including minimum, 1st quartile, median, mean, 3rd quartile, maximum, and number of missing values (NA). Plots are drawn to the current graphics device.

### Examples

```
dataBeta$CV <- sqrt(dataBeta$vardir)/dataBeta$y
explore(y~X1+X2, CV = "CV", data = dataBeta)
```

# Index

## \* **datasets**

dataBeta, [4](#)

dataBetaNS, [5](#)

betaTF, [2](#)

dataBeta, [4](#), [5](#)

dataBetaNS, [5](#)

explore, [6](#)

saeHB.TF.beta (saeHB.TF.beta-package), [2](#)

saeHB.TF.beta-package, [2](#)