

Package ‘saeczi’

July 23, 2025

Type Package

Title Small Area Estimation for Continuous Zero Inflated Data

Version 0.2.0

Maintainer Josh Yamamoto <joshuayamamoto5@gmail.com>

Description Provides functionality to fit a zero-inflated estimator for small area estimation.

This estimator is a combines a linear mixed effects regression model and a logistic mixed effects regression model via a two-stage modeling approach. The estimator's mean squared error is estimated via a parametric bootstrap method. Chandra and others (2012, <[doi:10.1080/03610918.2011.598991](https://doi.org/10.1080/03610918.2011.598991)>) introduce and describe this estimator and mean squared error estimator. White and others (2024+, <[doi:10.48550/arXiv.2402.03263](https://doi.org/10.48550/arXiv.2402.03263)>) describe the applicability of this estimator to estimation of forest attributes and further assess the estimator's properties.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports dplyr, lme4, purrr, progressr, furrr, future, rlang, Rcpp

RoxygenNote 7.3.1

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Depends R (>= 4.1.0)

LinkingTo Rcpp, RcppEigen

URL <https://harvard-ufds.github.io/saeczi/>

NeedsCompilation yes

Author Josh Yamamoto [aut, cre],

Dinan Elsyad [aut],

Grayson White [aut],

Julian Schmitt [aut],

Niels Korsgaard [aut],

Kelly McConville [aut],

Kate Hu [aut]

Repository CRAN
Date/Publication 2024-06-06 14:50:02 UTC

Contents

pop	2
saeczi	2
samp	4
Index	5

pop	<i>FIA Population Level Auxiliary Data for Oregon</i>
-----	---

Description

FIA Population Level Auxiliary Data for Oregon

Usage

pop

Format

An object of class `data.frame` with 10060 rows and 10 columns.

saeczi	<i>Fit a zero-inflation estimator.</i>
--------	--

Description

Fit a zero-inflation estimator.

Usage

```
saeczi(  
  samp_dat,  
  pop_dat,  
  lin_formula,  
  log_formula = lin_formula,  
  domain_level,  
  B = 100L,  
  mse_est = FALSE,  
  estimand = "means",  
  parallel = FALSE  
)
```

Arguments

samp_dat	A data.frame with domains, auxiliary variables, and the response variable of a sample
pop_dat	A data.frame with domains and auxiliary variables of a population.
lin_formula	Formula. Specification of the response and fixed effects of the linear regression model
log_formula	Formula. Specification of the response and fixed effects of the logistic regression model
domain_level	String. The column name in samp_dat and pop_dat that encodes the domain level
B	Integer. The number of bootstraps to be used in MSE estimation.
mse_est	Logical. Whether or not MSE estimation should happen.
estimand	String. Whether the estimates should be 'totals' or 'means'.
parallel	Logical. Should the MSE estimation be computed in parallel

Value

An object of class 'zi_mod' with defined 'print()' and 'summary()' methods. The object is structured like a list and contains the following elements:

- * call: The original function call
- * res: A data.frame containing the estimates and mse estimates
- * lin_mod: The modeling object used to fit the original linear model
- * log_mod: The modeling object used to fit the original logistic model

Examples

```
data(pop)
data(samp)

lin_formula <- DRYBIO_AG_TPA_live_ADJ ~ tcc16 + elev

result <- saeczi(samp_dat = samp,
                 pop_dat = pop,
                 lin_formula = lin_formula,
                 log_formula = lin_formula,
                 domain_level = "COUNTYFIPS",
                 mse_est = FALSE)
```

`samp`

*FIA sample data for Oregon***Description**

FIA sample data for Oregon

Usage

`samp`

Format

An object of class `tbl_df` (inherits from `tbl`, `data.frame`) with 1494 rows and 11 columns.

Index

* datasets

pop, [2](#)
samp, [4](#)

pop, [2](#)

saeczi, [2](#)
samp, [4](#)