

Package ‘bootcc’

August 28, 2022

Type Package

Title Bootstrap Inference for Logistic Regression in Case-Cohort Studies

Version 1.1-1

Date 2022-08-21

Author Hisashi Noma

Maintainer Hisashi Noma <noma@ism.ac.jp>

Description The conventional robust variance estimator for logistic regression in case-cohort studies is biased. To provide an unbiased estimate of standard error, the duplications of case and sub-cohort samples should be adequately accounted. This package provides bootstrap standard error estimates that accounted the duplications appropriately, and the corresponding confidence intervals and Wald-type p-values.

Depends R (>= 3.5.0)

Imports stats, MASS

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

R topics documented:

bootcc-package	1
bootcc	2
exdata	3

Index	4
--------------	----------

bootcc-package	<i>The ‘bootcc’ package.</i>
----------------	------------------------------

Description

Bootstrap Inference for Logistic Regression in Case-Cohort Studies.

References

Noma, H. (2022+). Variance estimation for logistic regression in case-cohort studies. Forthcoming.

bootcc

*Bootstrap Inference for Logistic Regression in Case-Cohort Studies***Description**

The conventional robust variance estimator for logistic regression in case-cohort studies is biased. To provide an unbiased estimate of standard error, the duplications of case and subcohort samples should be adequately accounted. This package provides bootstrap standard error estimates that accounted the duplications appropriately, and the corresponding confidence intervals and Wald-type p-values.

Usage

```
bootcc(formula, cases, subcohort, B=2000)
```

Arguments

formula	Formula of the logistic regression. Please use d for the outcome variable (=0,1).
cases	Dataset of the case samples. Please use y for the case/non-case status (=0,1).
subcohort	Dataset of the subcohort samples. Please use y for the case/non-case status (=0,1).
B	Number of the bootstrap resampling (default:2000).

Value

- coef: Regression coefficient estimates.
- SE: Bootstrap SE estimates.
- 95%CL: Lower limit of the 95% confidence intervals.
- 95%CU: Upper limit of the 95% confidence intervals.
- P-value: P-value for the tests of regression coefficients.

References

Noma, H. (2022+). Variance estimation for logistic regression in case-cohort studies. Forthcoming.

Examples

```
data(exdata)

cases <- exdata$cases # dataset of the case samples
subcohort <- exdata$subcohort # dataset of the subcohort samples

cases$d <- 1 # outcome variable = 1 for the case samples
subcohort$d <- 0 # outcome variable = 0 for the subcohort samples

cc <- rbind(cases,subcohort)

g1 <- glm(d ~ z + x1 + x2 ,data=cc,family=binomial(link="logit"))
summary(g1) # ordinary logistic regression; the SE estimate is biased.

bootcc(d ~ z + x1 + x2, cases=cases, subcohort=subcohort)
# the bootstrap inference for logistic regression
```

`exdata`*An example data of a case-cohort study*

Description

- `cases`: Dataset for case samples
- `subcohort`: Dataset for subcohort samples
- `id`: The subject ID
- `y`: An indicator variable that specify the case/non-case status
- `x1`: A binary covariate
- `x2`: A binary covariate
- `x3`: A binary covariate
- `z`: A binary covariate
- `z1`: A continous covariate

Usage`data(exdata)`**Format**

A list involving case and subcohort samples data

References

Noma, H. (2022+). Variance estimation for logistic regression in case-cohort studies. Forthcoming.

Index

* **datasets**

exdata, [3](#)

bootcc, [2](#)

bootcc-package, [1](#)

exdata, [3](#)