

# Package ‘ClusterBootstrap’

June 26, 2018

**Title** Analyze Clustered Data with Generalized Linear Models using the Cluster Bootstrap

**Date** 2018-06-26

**Version** 1.0.0

**Description** Provides functionality for the analysis of clustered data using the cluster bootstrap.

**Depends** R (>= 3.0), stats, utils, graphics, parallel

**License** GPL-3

**URL** <https://github.com/mathijsdeen/ClusterBootstrap>

**BugReports** <https://github.com/mathijsdeen/ClusterBootstrap/issues>

**Maintainer** Mathijs Deen <m.l.deen@fsw.leidenuniv.nl>

**LazyData** true

**RoxygenNote** 6.0.1

**Author** Mathijs Deen [aut, cre],  
Mark de Rooij [aut]

## R topics documented:

clusbootglm . . . . .	2
clusbootsample . . . . .	3
coef.clusbootglm . . . . .	4
confint.clusbootglm . . . . .	4
medication . . . . .	5
opposites . . . . .	6
summary.clusbootglm . . . . .	6
<b>Index</b>	<b>8</b>

clusbootglm

*Fit generalized linear models with the cluster bootstrap***Description**

Fit a generalized linear model with the cluster bootstrap for analysis of clustered data.

**Usage**

```
clusbootglm(model, data, clusterid, family = gaussian, B = 5000,
  confint.level = 0.95, n.cores = 1)
```

**Arguments**

model	generalized linear model to be fitted with the cluster bootstrap.
data	dataframe that contains the data.
clusterid	variable in data that identifies the clusters.
family	error distribution to be used in the model, e.g. gaussian or binomial.
B	number of bootstrap samples.
confint.level	level of confidence interval.
n.cores	number of CPU cores to be used.

**Details**

Some useful methods for the obtained clusbootglm class object are [summary.clusbootglm](#), [coef.clusbootglm](#), and [clusbootsample](#).

**Value**

clusbootglm produces an object of class "clusbootglm", containing the following relevant components:

coefficients	A matrix of B rows, containing the parameter estimates for all bootstrap samples.
bootstrap.matrix	n*B matrix, of which each column represents a bootstrap sample; each value in a column represents a unit of subjectid.
lm.coefs	Parameter estimates from a single (generalized) linear model.
boot.coefs	Mean values of the parameter estimates, derived from the bootstrap coefficients.
boot.sds	Standard deviations of cluster bootstrap parameter estimates.
ci.level	User defined confidence interval level.
percentile.interval	Confidence interval based on percentiles, given the user defined confidence interval level.
parametric.interval	Confidence interval based on lm.coefs and column standard deviations of coefficients, given the user defined confidence interval level.
BCa.interval	Confidence interval based on percentiles with bias correction and acceleration, given the user defined confidence interval level.

`samples.with.NA.coef`

Cluster bootstrap sample numbers with at least one coefficient being NA.

`failed.bootstrap.samples`

For each of the coefficients, the number of failed bootstrap samples are given.

### Author(s)

Mathijs Deen, Mark de Rooij

### Examples

```
## Not run:
data(opposites)
clusbootglm(SCORE~Time*COG,data=opposites,clusterid=Subject)
## End(Not run)
```

---

`clusbootstrap`

*Return data for specified bootstrap sample*

---

### Description

Returns the full data frame for a specified bootstrap sample in a `clusbootglm` object.

### Usage

```
clusbootstrap(object, samplenr)
```

### Arguments

<code>object</code>	object of class <code>clusbootglm</code> , created with the <code>clusbootglm</code> function.
<code>samplenr</code>	sample number for which the data frame should be returned.

### Author(s)

Mark de Rooij, Mathijs Deen

### Examples

```
## Not run:
data(opposites)
cbglm.1 <- clusbootglm(SCORE~Time*COG,data=opposites,clusterid=Subject)
clusbootstrap(cbglm.1, samplenr=1)
## End(Not run)
```

---

coef.clusbootglm	<i>Obtain coefficients from cluster bootstrap object</i>
------------------	--

---

### Description

Returns the coefficients of an object of class clusbootglm.

### Usage

```
## S3 method for class 'clusbootglm'
coef(object, estimate.type = "bootstrap", ...)
```

### Arguments

object	object of class clusbootglm.
estimate.type	type of coefficient (bootstrap or GLM).
...	other arguments.

### Author(s)

Mathijs Deen

### Examples

```
## Not run:
data(opposites)
cbglm.1 <- clusbootglm(SCORE~Time*COG,data=opposites,clusterid=Subject)
coef(cbglm.1, estimate.type="bootstrap")
## End(Not run)
```

---

confint.clusbootglm	<i>Confidence intervals for cluster bootstrap model parameters</i>
---------------------	--

---

### Description

Computes confidence intervals for one or more parameters in a fitted GLM with the cluster bootstrap.

### Usage

```
## S3 method for class 'clusbootglm'
confint(object, parm = "all", level = 0.95,
  interval.type = "BCa", ...)
```

**Arguments**

<code>object</code>	object of class <code>clusbootglm</code> .
<code>parm</code>	a specification of which parameters are to be given confidence intervals, either a vector of numbers or a vector of names. Defaults to all parameters.
<code>level</code>	the required confidence level
<code>interval.type</code>	type of confidence level. Options are BCa, percentile, and parametric.
<code>...</code>	other arguments.

**Author(s)**

Mathijs Deen

**Examples**

```
## Not run:
data(opposites)
cbglm.1 <- clusbootglm(SCORE~Time*COG,data=opposites,clusterid=Subject)
confint(cbglm.1,parm=c("Time","COG"), level=.90, interval.type="percentile")
## End(Not run)
```

---

medication	<i>Medication data</i>
------------	------------------------

---

**Description**

The medication dataframe consists of 1242 observations within 73 individuals that were part of a placebo controlled clinical trial, as reported in Tomarken, Shelton, Elkins, and Anderson (1997).

The data were retrieved from the accompanied website of Singer & Willett (2003), at <https://stats.idre.ucla.edu/other/exam>

**Usage**

```
medication
```

**Format**

the following variables are available:

- `id`: subject indicator
- `treatment`: either placebo (0) or antidepressant (1)
- `time`: number of days since trial start.
- `pos`: positive affect. Higher scores indicate a more positive mood.

**References**

- Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. NY: Oxford University Press, Inc.
- Tomarken, A.J., Shelton, R.C., Elkins, L., & Anderson, T (1997). *Sleep deprivation and anti-depressant medication: Unique effects on positive and negative affect*. Poster session presented at the 9th annual meeting of the American Psychological Society, Washington, DC.

---

opposites

*Opposites naming data*

---

### Description

The opposites dataframe consists of 144 observations within 36 individuals that completed an inventory that assesses their performance on a timed cognitive task called "opposites naming".

The dataset does not contain the empirical data within 35 individuals from the experiment by Willett (1988), but a simulation based on the multilevel model from Singer & Willett (2003) within 36 individuals.

### Usage

opposites

### Format

the following variables are available:

- Subject: subject indicator
- Time: a time variable, ranging 0-3
- COG: cognitive skill, measured once (at time=0)
- SCORE: score on opposites naming task

### References

- Willett, J.B. (1988). Questions and answers in the measurement of change. In: E. Rothkopf (Ed.), *Review of research in education (1988-89)* (pp. 345-422). Washington, DC: American Educational Research Association.
- Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. NY: Oxford University Press, Inc.

---

summary.clusbootglm

*Summarize output of cluster bootstrap GLM*

---

### Description

Returns the summary of an object of class clusbootglm.

### Usage

```
## S3 method for class 'clusbootglm'
summary(object, estimate.type = "bootstrap",
        interval.type = "BCa", ...)
```

**Arguments**

object	object of class clusbootglm.
estimate.type	specify which type of estimate should be returned, either bootstrap means (default) or GLM estimates from model fitted on original data.
interval.type	which confidence interval should be used. Options are parametric, percentile, and BCa intervals.
...	other arguments.

**Author(s)**

Mathijs Deen

**Examples**

```
## Not run:  
data(opposites)  
cbglm.1 <- clusbootglm(SCORE~Time*COG,data=opposites,clusterid=Subject)  
summary(cbglm.1, interval.type="percentile")  
## End(Not run)
```

# Index

`clusbootglm`, [2](#)  
`clusbootsample`, [2](#), [3](#)  
`coef.clusbootglm`, [2](#), [4](#)  
`confint.clusbootglm`, [4](#)  
  
`medication`, [5](#)  
  
`opposites`, [6](#)  
  
`summary.clusbootglm`, [2](#), [6](#)