

Package ‘RDStreeboot’

November 24, 2016

Type Package

Title RDS Tree Bootstrap Method

Version 1.0

Date 2016-11-23

Author Aaron J. Baraff

Maintainer Aaron J. Baraff <ajbaraff@uw.edu>

Description A tree bootstrap method for estimating uncertainty in respondent-driven samples (RDS). Quantiles are estimated by multilevel resampling in such a way that preserves the dependencies of and accounts for the high variability of the RDS process.

License GPL-2 | GPL-3

NeedsCompilation no

Repository CRAN

Date/Publication 2016-11-24 00:36:48

R topics documented:

RDStreeboot-package	1
faux.network	2
sample.RDS	3
treeboot.RDS	5

Index	7
--------------	----------

RDStreeboot-package *RDStreeboot: Tree Bootstrap Method*

Description

A tree bootstrap method for estimating uncertainty in respondent-driven samples (RDS). Quantiles are estimated by multilevel resampling in such a way that preserves the dependencies of and accounts for the high variability of the RDS process.

Details

Package: RDStreeboot
Type: Package
Version: 1.0
Date: 2016-11-23
License: GPL-2 | GPL-3

The main estimation function is `treeboot.RDS`. It produces estimates for the quantiles of traits from a respondent-driven sample (RDS) using the tree bootstrap method. Options allow for different quantiles to be estimated and the number of tree bootstrap samples to be drawn. Also included is a function to draw a random respondent-driven sample (RDS) from a social network (`sample.RDS`). A faux social network dataset has been provided for testing and examples (`faux.network`).

Author(s)

Aaron J. Baraff

Maintainer: Aaron J. Baraff <ajbaraff@uw.edu>

References

Baraff, A. J., McCormick, T. H., and Raftery, A. E., "Estimating uncertainty in respondent-driven sampling using a tree bootstrap method."

Examples

```
## load data
data(faux.network)

## draw RDS from network
samp <- sample.RDS(faux.network$traits, faux.network$adj.mat, 100, 2, 3, c(0,1/3,1/3,1/3), TRUE)

## estimate 80% and 95% confidence intervals
treeboot.RDS(samp, c(0.025, 0.10, 0.90, 0.975), 2000)
```

faux.network

Faux Social Network Dataset

Description

This dataset contains a faux social network created for testing and examples.

Usage

```
data("faux.network")
```

Format

A list with two components:

traits a data frame, the first column giving ID numbers for the nodes in the social network and the remaining columns giving continuous or discrete (not categorical) traits for each node.

adj.mat a matrix of 0's and 1's, the adjacency matrix for the social network.

Details

This faux social network contains 1000 nodes and 5000 edges with each node having three binary traits labeled "X", "Y", and "Z".

See Also

[sample.RDS](#)

Examples

```
## load data
data(faux.network)

## draw RDS from network
samp <- sample.RDS(faux.network$traits, faux.network$adj.mat, 100, 2, 3, c(0,1/3,1/3,1/3), TRUE)
```

sample.RDS

Draw RDS Sample

Description

This function draws a respondent-driven sample (RDS) from a social network.

Usage

```
sample.RDS(traits, adj.mat, n, num.seeds, num.samp, num.prob, replace)
```

Arguments

traits	a data frame, the first column giving ID numbers for the nodes in the social network and the remaining columns giving continuous or discrete (not categorical) traits for each node.
adj.mat	a matrix of 0's and 1's, the adjacency matrix for the social network.
n	a positive integer, the size of the sample to be drawn, with default 100.
num.seeds	a positive integer, the number of seeds from which the RDS begins, with default 1.
num.samp	a positive integer, the maximum number of recruits for each respondent, with default 2.

num.prob	a vector of positive numbers with length <code>num.samp + 1</code> which sum to 1, the probability of each respondent recruiting 0, 1, ..., <code>num.samp</code> new respondents, with default <code>c(0, ..., 1)</code> .
replace	a boolean, whether sampling from the social network is performed with replacement, with default <code>TRUE</code> .

Details

The function `sample.RDS` draws a random respondent-driven sample (RDS) from a social network. First, seeds are drawn with probability proportional to degree. Then each respondent successively recruits additional respondents until the desired sample size is attained. Options allow for different numbers of seeds, the number and probability distribution of recruits, and whether or not sampling is done with replacement. Each call to the function results in a single RDS realization.

Value

A list with four components:

nodes	a vector of numbers, the ID numbers of the drawn nodes in the social network.
edges	a list of two vectors of positive integers, the edge list of the RDS. For each edge, <code>node1</code> lists the respondent (source) nodes and <code>node2</code> lists the recruited (destination) nodes.
degree	a vector of positive integers, the degree of each sampled node in the social network.
traits	a data frame, the columns giving continuous or discrete (not categorical) traits for each node.

Author(s)

Aaron J. Baraff

Maintainer: Aaron J. Baraff <ajbaraff at uw.edu>

References

Baraff, A. J., McCormick, T. H., and Raftery, A. E., "Estimating uncertainty in respondent-driven sampling using a tree bootstrap method."

Examples

```
## load data
data(faux.network)

## draw RDS from network
samp <- sample.RDS(faux.network$traits, faux.network$adj.mat, 100, 2, 3, c(0,1/3,1/3,1/3), TRUE)
```

`treeboot.RDS`*Estimate RDS Quantiles*

Description

This function estimates quantiles of numerical traits from a respondent-driven sample (RDS) using the tree bootstrap method.

Usage

```
treeboot.RDS(samp, quant, B)
```

Arguments

<code>samp</code>	A list with four components: <code>nodes</code> a vector of numbers, the ID numbers of the sampled nodes. <code>edges</code> a list of two vectors of positive integers, the edge list of the RDS. For each edge, <code>node1</code> lists the respondent (source) nodes and <code>node2</code> lists the recruited (destination) nodes. <code>degree</code> a vector of positive integers, the degree of each sampled node in the underlying social network. <code>traits</code> a data frame, the columns giving continuous or discrete (not categorical) traits for each node.
<code>quant</code>	a vector of numbers between 0 and 1, the quantiles to be estimated for each trait, with default <code>c(0.025, 0.10, 0.90, 0.975)</code> .
<code>B</code>	a positive integer, the number of tree bootstrap samples to be drawn, with default 2000.

Details

The function `treeboot.RDS` estimates the quantiles of traits from a respondent-driven sample (RDS) using the tree bootstrap method. Options allow for different quantiles to be estimated and the number of tree bootstrap samples to be drawn.

Value

A matrix with one row for each trait in the RDS. Each row has a single computed quantile corresponding to the input vector `quant`.

Author(s)

Aaron J. Baraff

Maintainer: Aaron J. Baraff <ajbaraff@uw.edu>

References

Baraff, A. J., McCormick, T. H., and Raftery, A. E., "Estimating uncertainty in respondent-driven sampling using a tree bootstrap method."

Examples

```
## load data
data(faux.network)

## draw RDS from network
samp <- sample.RDS(faux.network$traits, faux.network$adj.mat, 100, 2, 3, c(0,1/3,1/3,1/3), TRUE)

## estimate 80% and 95% confidence intervals
treeboot.RDS(samp, c(0.025, 0.10, 0.90, 0.975), 2000)
```

Index

*Topic **datasets**

[faux.network](#), [2](#)

*Topic **package**

[RDStreeboot-package](#), [1](#)

[faux.network](#), [2](#), [2](#)

[RDStreeboot \(RDStreeboot-package\)](#), [1](#)

[RDStreeboot-package](#), [1](#)

[sample.RDS](#), [2](#), [3](#), [3](#)

[treeboot.RDS](#), [2](#), [5](#)