**Package ‘bootImpute’**

July 18, 2019

**Type**  Package

**Title**  Bootstrap Inference for Multiple Imputation

**Version**  1.0.0

**Author**  Jonathan Bartlett

**Maintainer**  Jonathan Bartlett <j.w.bartlett@bath.ac.uk>

**Description**  Bootstraps and imputes incomplete datasets. Then performs inference on estimates obtained from analysing the imputed datasets as proposed by von Hippel (2018) <arXiv:1210.0870v9>.

**Depends**  R (>= 2.10)

**License**  GPL-3

**Encoding**  UTF-8

**LazyData**  true

**RoxygenNote**  6.1.1

**Imports**  mice, smcfcs

**Suggests**  testthat

**NeedsCompilation**  no

**Repository**  CRAN

**Date/Publication**  2019-07-18 06:36:34 UTC

---

**R topics documented:**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>bootImpute</td>
<td>1</td>
</tr>
<tr>
<td>bootImputeAnalyse</td>
<td>3</td>
</tr>
<tr>
<td>bootMice</td>
<td>4</td>
</tr>
<tr>
<td>bootSmcfcs</td>
<td>4</td>
</tr>
<tr>
<td>ex_linquad</td>
<td>5</td>
</tr>
</tbody>
</table>
**bootImpute**

**Bootstrap then impute an incomplete dataset**

Description

Bootsraps an incomplete dataset and then imputes each bootstrap a number of times. The resulting list of bootstrapped then imputed datasets can be analysed with `bootImputeAnalyse`.

Usage

```
bootImpute(obsdata, impfun, nBoot = 200, nImp = 2, ...)
```

Arguments

- `obsdata` : The data frame to be imputed.
- `impfun` : A function which when passed an incomplete dataset will return a single imputed data frame.
- `nBoot` : The number of bootstrap samples to take. It is recommended that you use a minimum of 200.
- `nImp` : The number of times to impute each bootstrap sample. Two is recommended.
- `...` : Other parameters that are to be passed through to `impfun`.

Details

The function can be used any kind of multiple imputation procedure. The `impfun` must be a function which when passed an incomplete datasets and possibly additional arguments, returns a single imputed data frame. Depending on what imputation function returns by default, you may need to write a small wrapper function that calls the imputation procedure once and returns the imputed dataset. See the Example for an illustration with the `mice` package.

Value

A list of imputed datasets.

Examples

```
#this example shows how you can use bootImpute to impute using the mice package. If you do want to impute using MICE you can instead use the #bootMice function, which essentially contains the code below
library(mice)

set.seed(564764)

#write a wrapper function to call mice with one imputation and return #the imputed dataset
impOnce <- function(inputData) {
  oneImp <- mice::mice(inputData, m=1)
}
bootImputeAnalyse

```r
mice::complete(oneImp)
}

#bootstrap twice and impute each twice
#in practice you should bootstrap many more times, e.g. at least 200
imps <- bootImpute(ex_linquad, impOnce, nBoot=2, nImp=2)
```
#write a wrapper to analyse an imputed dataset
analyseImp <- function(inputData) {
  coef(lm(y~z+x+xsq, data=inputData))
}
est <- bootImputeAnalyse(imps, analyseImp)

---

**bootMice**

*Bootstrap then impute using mice*

**Description**

Bootstrap an incomplete dataset and then impute each bootstrap a number of times using the mice package. The resulting list of bootstrapped then imputed datasets can be analysed with bootImputeAnalyse. To run this function requires the mice package to be installed.

**Usage**

`bootMice(obsdata, nBoot = 200, nImp = 2, ...)`

**Arguments**

- `obsdata` The data frame to be imputed.
- `nBoot` The number of bootstrap samples to take. It is recommended that you use a minimum of 200.
- `nImp` The number of times to impute each bootstrap sample. Two is recommended.
- `...` Other arguments that are to be passed to mice.

**Value**

A list of imputed datasets.

**Examples**

```r
library(mice)
set.seed(564764)
head(ex_linquad)
#bootstrap 10 times and impute each twice
imps <- bootMice(ex_linquad, nBoot=10, nImp=2)
```
Description

Bootstrap then impute using smcfcs

Bootstraps an incomplete dataset and then imputes each bootstrap a number of times using the smcfcs package. The resulting list of bootstrapped then imputed datasets can be analysed with bootImputeAnalyse. To run this function requires the smcfcs package to be installed.

Usage

bootSmcfcs(obsdata, nBoot = 200, nImp = 2, ...)

Arguments

obsdata  The data frame to be imputed.
nBoot  The number of bootstrap samples to take. It is recommended that you use a minimum of 200.
nImp  The number of times to impute each bootstrap sample. Two is recommended.
...  Other arguments that are to be passed to smcfcs.

Value

A list of imputed datasets.

Examples

library(smcfcs)
set.seed(564764)
head(ex_linquad)
#bootstrap twice and impute each twice
#in practice you should bootstrap many more times, e.g. at least 200
imps <- bootSmcfcs(ex_linquad, nBoot=2, nImp=2,
    smtype="lm", smformula="y~z+x+xsq",
    method=c("","","norm","x^2",""))

ex_linquad  Simulated example data with continuous outcome and quadratic covariate effects

Description

A dataset containing simulated data where the outcome depends quadratically on a partially observed covariate.
Usage

`ex_linquad`

Format

A data frame with 1000 rows and 5 variables:

- **y** Continuous outcome
- **z** Fully observed covariate, with linear effect on outcome
- **x** Partially observed normally distributed covariate, with quadratic effect on outcome
- **xsq** The square of x, which thus has missing values also
- **v** An auxiliary variable (i.e. not contained in the substantive model)