Package 'simputation'

July 22, 2016

Maintainer Mark van der Loo <mark.vanderloo@gmail.com>

License GPL-3

Title Simple Imputation

LazyData no

Type Package

LazyLoad yes

Description
Easy to use interfaces to a number of imputation methods that fit in the not-a-pipe operator of the 'magrittr' package.

Version 0.0.0.1

Imports stats, MASS

URL https://github.com/markvanderloo/simputation

BugReports https://github.com/markvanderloo/simputation/issues

Date 2016-07-21

Suggests testthat

RoxygenNote 5.0.1

R topics documented:

`impute_` *Impute missing data*

Description

Use to fit and impute missing data.

Usage

```
impute_lm(data, model, add_residual = c("none", "observed", "normal"), ...)

impute_rlm(data, model, add_residual = c("none", "observed", "normal"), ...)

impute_const(data, model, add_residual = c("none", "observed", "normal"), ...)

impute_median(data, model, add_residual = c("none", "observed", "normal"),
  ...)

impute_proxy(data, model, ...)
```

Arguments

`data`	The data
`model`	`[formula]` imputation model description (see Details below).
`add_residual`	`[character]` Type of residual to add. `"normal"` means that the imputed value is drawn from `N(mu,sd)` where `mu` and `sd` are estimated from the model's residuals (`mu` should equal zero in most cases). If `add_residual = "observed"`, residuals are drawn (with replacement) from the model's residuals.
`...`	further arguments passed to `lm` or `rlm`

Value

`data`, imputed according to `model`.

Details

Model specification works as usual, except that it is possible to impute multiple variables based on the same model. To specify the same model for multiple variables, simply add variables to the left-hand side of the formula using `+`. Also see the examples.

If a value cannot be imputed because one of its predictors is missing, the value will remain missing after imputation.

If a model cannot be fitted, e.g. because the imputed model is missing, a warning is emitted and fot htat variable no imputation will take place.

Model descriptions

Model	description
impute_lm	Use stats::lm to train the imputation model.
impute_rlm	Use MASS::rlm to train the imputation model.
impute_median	Median imputation. Predictors are treated as grouping variables for computing medians.
impute_const	Impute a constant value
impute_proxy	Copy a value from the predictor variable.

Examples

```
data(iris)
irisNA <- iris
irisNA[1:4, "Sepal.Length"] <- NA
irisNA[3:7, "Sepal.Width"] <- NA

# impute a single variable (Sepal.Length)
i1 <- impute_lm(irisNA, Sepal.Length ~ Sepal.Width + Species)

# impute both Sepal.Length and Sepal.Width, using robust linear regression
i2 <- impute_rlm(irisNA, Sepal.Length + Sepal.Width ~ Species + Petal.Length)
```

---

simputation *simputation*

Description

simputation

Index