Package 'stddiff'

June 10, 2022

Type Package

Title Calculate the Standardized Difference for Numeric, Binary and Category Variables

Version 3.1

Author Zhicheng Du, Yuantao Hao

Maintainer Zhicheng Du<dgdzc@hotmail.com>

Description Contains three main functions including stddiff.numeric(), stddiff.binary() and stddiff.category(). These are used to calculate the standardized difference between two groups. It is especially used to evaluate the balance between two groups before and after propensity score matching.

License GPL-3

NeedsCompilation no

Repository CRAN

Date/Publication 2022-06-10 09:40:02 UTC

R topics documented:

`stddiff` *Calculate the Standardized Difference for Numeric, Binary and Category Variables*

Description

Contains three main functions including stddiff.numeric(), stddiff.binary() and stddiff.category(). These are used to calculate the standardized difference between two groups. It is especially used to evaluate the balance between two groups before and after propensity score matching.

Usage

```
stddiff.numeric(data,gcol,vcol)
stddiff.binary(data,gcol,vcol)
stddiff.category(data,gcol,vcol)
```

Arguments

data	a dataframe
gcol	a column number of group variable in data, 0 for control group, 1 for treatment group
vcol	one or more column numbers of different types variables in data

Details

stddiff.numeric() is used for the numeric variables. For the skewed variables, you should change to the rank using the rank() function before computing the "stddiff".

stddiff.binary() is used for the binomial variables.

stddiff.category() is used for the categorical variables.

Imbalance was usually defined as "stddiff" greater than 0.1 or 0.2 (which means the small effect size).

Value

for stddiff.numeric function:

mean.c	the mean of control group
sd.c	the standard deviation of control group
mean.t	the mean of treatment group
sd.t	the standard deviation of treatment group
missing.c	the counts of missing value of control group
missing.t	the counts of missing value of treatment group stddiff: the standardized difference between two groups
stddiff.l	the lower limit of the 95 percentage confidence interval of standardized difference
stddiff.u	the upper limit of the 95 percentage confidence interval of standardized difference

for stddiff.binary function:

p.c	the proportion of last level in the control group
p.t	the proportion of last level in the treatment group
missing.c	the counts of missing value of control group
missing.t	the counts of missing value of treatment group
stddiff	the standardized difference between two groups

stddiff.l	the lower limit of the 95 percentage confidence interval of standardized difference
stddiff.u	the upper limit of the 95 percentage confidence interval of standardized difference

for stddiff.category function:

p.c	the proportion of each level in the control group
p.t	the proportion of each level in the treatment group
missing.c	the counts of missing value of control group
missing.t	the counts of missing value of treatment group
stddiff	the standardized difference between two groups
stddiff.l	the lower limit of the 95 percentage confidence interval of standardized difference
stddiff.u	the upper limit of the 95 percentage confidence interval of standardized difference

Note

Update:

version 2.0: Avoiding the negative number for the 'stddiff' of stddiff.numeric() and stddiff.binary()

version 3.0: Fixing the incorrect format in the results of stddiff.category()

version 3.1: Fixing the incorrect counts of missing values of stddiff.numeric(), stddiff.binary(), stddiff.category()

Author(s)

Zhicheng Du<dgdzc@hotmail.com>, Yuantao Hao<haoyt@mail.sysu.edu.cn>

References

Yang DS, Dalton JE. A Unified Approach to Measuring the Effect Size Between Two Groups Using SAS. SAS Global Forum 2012. paper 335

See Also

nothing

Examples

```
set.seed(2016)
treat<-round(abs(rnorm(100)+1)*10,0)
numeric<-round(abs(rnorm(100)+1)*10,0)
binary<-round(abs(rnorm(100)+1)*10,0)
category<-round(abs(rnorm(100)+1)*10,0)
data<-data.frame(treat,numeric,binary,category)
stddiff.numeric(data=data,gcol=1,vcol=c(2,2))
#stddiff.binary(data=data,gcol=1,vcol=c(3,3))
#stddiff.category(data=data,gcol=1,vcol=c(4,4))
```

Index