Package 'smd'

May 6, 2024

Type Package

Title Compute Standardized Mean Differences

Version 0.7.0

Description Computes standardized mean differences and confidence intervals for multiple data types based on Yang, D., & Dalton, J. E. (2012) <https://support.sas.com/resources/papers/proceedings12/335-2012.pdf>.

Imports MASS (>= 7.3-50), methods (>= 3.5.1)

Suggests testthat, stddiff, tableone, knitr, dplyr, purrr, markdown, rmarkdown

License MIT + file LICENSE

URL https://bsaul.github.io/smd/

BugReports https://github.com/bsaul/smd/issues

Encoding UTF-8

RoxygenNote 7.3.1

VignetteBuilder knitr

Repository CRAN

NeedsCompilation no

Author Bradley Saul [aut, cre],
Alex Breskin [ctb],
Catie Wiener [ctb],
Matt Phelan [ctb],
Daniel Sjoberg [ctb],
Nuvan Rathnayaka [ctb]

Maintainer Bradley Saul <bradleysaul@fastmail.com>

Date/Publication 2024-05-06 19:30:02 UTC

R topics documented:

smd *Compute Standardized Mean Difference*

Description

Computes the standardized mean differnce (SMD) between two groups.

$$d = \sqrt{D' S^{-1} D}$$

where D is a vector of differences between group 1 and 2 and S is the covariance matrix of these differences. If D is length 1, the result is multplied by $sign(D)$.

In the case of a `numeric` or `integer` variable, this is equivalent to:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2 + s_2^2)/2}}$$

where $\bar{x}_g$ is the sample mean for group g and s_g^2 is the sample variance.

For a `logical` or `factor` with only two levels, the equation above is $\bar{x}_g = \hat{p}_g$, i.e. the sample proportion and $s_g^2 = \hat{p}_g(1 - \hat{p}_g)$.

When using the SMD to evaluate the effectiveness of weighting in achieving covariate balance, it is important to isolate the change in SMD before and after weighting to the change in mean difference, so the denominator (covariance matrix) must be held constant (Stuart 2008, doi:10.1002/sim.3207). By default, the unweighted covariance matrix is used to compute SMD in both the unweighted and weighted case. If the weights are not being used to adjust for covariate imbalance (e.g. case weights), the `unwgt.var` argument can be set to `FALSE` to use the weighted covariance matrix as the denominator.

Usage

```
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'character,ANY,missing'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'character,ANY,numeric'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'logical,ANY,missing'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'logical,ANY,numeric'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'matrix,ANY,missing'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)
```

```
## S4 method for signature 'matrix,ANY,numeric'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'list,ANY,missing'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'list,ANY,numeric'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'data.frame,ANY,missing'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)

## S4 method for signature 'data.frame,ANY,numeric'
smd(x, g, w, std.error = FALSE, na.rm = FALSE, gref = 1L, unwgt.var = TRUE)
```

Arguments

x	a vector or matrix of values
g	a vector of at least 2 groups to compare. This should coercable to a factor.
w	a vector of numeric weights (optional)
std.error	Logical indicator for computing standard errors using compute_smd_var. Defaults to FALSE.
na.rm	Remove NA values from x? Defaults to FALSE.
gref	an integer indicating which level of g to use as the reference group. Defaults to 1.
unwgt.var	Use unweighted or weighted covariance matrix. Defaults to TRUE

Value

a data.frame containing standardized mean differences between levels of g for values of x. The data.frame contains the columns:

- term: the level being comparing to the reference level
- estimate: SMD estimates
- std.error: (if std.error = TRUE) SMD standard error estimates

Examples

```
x <- rnorm(100)
g <- rep(1:2, each = 50)
smd(x, g)
```

Index