Package ‘Partiallyoverlapping’

December 13, 2018

Type Package
Title Partially Overlapping Samples Tests
Version 2.0
Date 2018-12-12
Author Ben Derrick
Maintainer Ben Derrick <ben.derrick@uwe.ac.uk>
Depends R (>= 3.1.3)
Description Tests for a comparison of two partially overlapping samples.
   A comparison of means using the partially overlapping samples t-test:
   See Derrick, Russ, Toher and White (2017), Test
   statistics for the comparison of means for two samples which include
   both paired observations and independent observations, Journal of
   Modern Applied Statistical Methods, 16(1).
   A comparison of proportions using the partially overlapping samples z-test:
   See Derrick, Dobson-Mckittrick, Toher and White (2015), Test statistics
   for comparing two proportions with partially overlapping samples.
   Journal of Applied Quantitative Methods, 10(3).
LazyData Yes
License GPL-3
Imports stats
RoxygenNote 6.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2018-12-13 14:00:12 UTC

R topics documented:

Partover.test ........................................................... 2
Prop.test ................................................................. 3

Index 6
The partially overlapping samples t-test

Description

Performs a comparison of means using the partially overlapping t-test, for two samples each with paired and unpaired observations. This function calculates the test statistic, the degrees of freedom, and the p-value. Additionally calculates a confidence interval for the difference in means when requested.

By default, four vectors are to be specified: unpaired observations in Sample 1, unpaired observations in Sample 2, paired observations in Sample 1, paired observations in Sample 2. If the structure of your data is of two vectors, one for each sample, then the option stacked = TRUE can be specified.

Usage

```
partover.test(x1 = NULL, x2 = NULL, x3 = NULL, x4 = NULL,
              var.equal = FALSE, mu = 0, alternative = "two.sided",
              conf.level = NULL, stacked = FALSE)
```

Arguments

- `x1`: a vector of unpaired observations in Sample 1 (or all observations in Sample 1 if stacked = "TRUE")
- `x2`: a vector of unpaired observations in Sample 2 (or all observations in Sample 2 if stacked = "TRUE")
- `x3`: a vector of paired observations in Sample 1 (not applicable if stacked = "TRUE")
- `x4`: a vector of paired observations in Sample 2 (not applicable if stacked = "TRUE")
- `var.equal`: a logical variable indicating whether to treat the two variances as being equal. If "TRUE" then the pooled variance is used to estimate the variance, otherwise the Welch approximation to the degrees of freedom is used.
- `mu`: difference in population means under the null hypothesis
- `alternative`: a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".
- `conf.level`: confidence level of the interval.
- `stacked`: indicator of whether paired and unpaired observations are stacked within one vector ("TRUE"), or if specified as four separate vectors (default)

Details

The formula is only applicable for the 2 sample partially overlapping samples t-test. The number of unpaired observations may be zero for up to one of the two samples. The number of paired observations must be of equal length of two or greater. Error messages are generated when these conditions are not true.

If your observations are from a data frame of two paired samples with missing observations, use the option "stacked=TRUE". Corresponding pairs should be given on the same row when this option is applied.
Value

A list which contains the following components of the test:

- statistic ~~ The value of the t-statistic
- parameter ~~ The degrees of freedom for the test statistic
- p.value ~~ The p-value for the test
- estimate ~~ The estimated difference in the means
- conf.int ~~ A confidence interval for the mean difference appropriate to the specified alternative hypothesis

Examples

```r
# Example taken from Derrick, Toher and White (2017)
# How to compare the means of two samples that include
# paired observations and independent observations:
# A companion to Derrick, Russ, Toher and White (2017).
# The Quantitative Methods in Psychology, 13(2). pp.120-126.

#The sample means for two groups, “a” and “b” are compared
#for a two sided test assuming equal variances.

#Approach 1:
#For each sample, unpaired observations and paired observations defined as separate vectors:
a.unpaired<-c(20,21,16,18,14,12,14,17)
a.paired<-c(14,15,18,20,11,19,14,15)
b.unpaired<-c(10,16,18,16,15,14,13,10)
b.paired<-c(15,10,15,17,13,19,12,13)
Partover.test(a.unpaired,b.unpaired,a.paired,b.paired,var.equal=TRUE)
# p.value = 0.026, the samples from group "a" and group "b" have significantly different means

#Equivalently, Approach 2:
#Independent observations and the paired samples stacked for each sample:
a<-c(20,21,16,18,14,12,14,17,NA,NA,NA,NA,NA,NA,NA,NA,14,15,18,20,11,19,14,15)
b<-c(NA,NA,NA,NA,NA,NA,NA,NA,10,16,18,16,15,14,13,10,15,10,15,17,13,19,12,13)
Partover.test(a,b,var.equal=TRUE,stacked=TRUE)
# p.value = 0.026, the samples from group "a" and group "b" have significantly different means
```

Prop.test

The partially overlapping samples z-test for dichotomous variables

Description

Performs a comparison of proportions using the partially overlapping z-test, for two dichotomous samples each with paired and unpaired observations. This functions calculates the test statistic, and the p-value. Additionally calculates a confidence interval for the difference in means when requested.
Prop.test

Usage

Prop.test(x1 = NULL, x2 = NULL, x3 = NULL, x4 = NULL, alternative = "two.sided", conf.level = NULL, stacked = FALSE)

Arguments

- **x1**: A vector of unpaired observations in Sample 1 (or all observations in Sample 1 if stacked = "TRUE")
- **x2**: A vector of unpaired observations in Sample 2 (or all observations in Sample 2 if stacked = "TRUE")
- **x3**: A vector of paired observations in Sample 1 (not applicable if stacked = "TRUE")
- **x4**: A vector of paired observations in Sample 2 (not applicable if stacked = "TRUE")
- **alternative**: A character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".
- **conf.level**: Confidence level of the interval.
- **stacked**: Indicator of whether paired and unpaired observations are stacked within one vector ("TRUE"), or if specified as four separate vectors (default)

Details

By default, four vectors are to be specified: unpaired observations in Sample 1, unpaired observations in Sample 2, paired observations in Sample 1, paired observations in Sample 2. If the structure of your data is of two vectors, one for each sample, then the option stacked = TRUE can be specified.

The formula is only applicable for the 2 sample partially overlapping samples t-test. The number of unpaired observations may be zero for up to one of the two samples. The number of paired observations must be of equal length of two or greater. Error messages are generated when these conditions are not true. Assumes the raw count data is in a data frame of "1"s and "0"s. Missing observations are recorded as "NA", for unpaired observations the corresponding observation for the "missing" pair is recorded as "NA". Corresponding pairs should be given on the same row when this option is applied.

Value

A list which contains the following components of the test:
- **statistic**: The value of the z-statistic
- **p.value**: The p-value for the test
- **estimate**: The estimated difference in proportions
- **conf.int**: A confidence interval for the difference in proportions appropriate to the specified alternative hypothesis

Examples

# Example taken from Derrick, Dobson-Mckittrick, Toher and White, (2015)
# Test statistics for comparing two proportions with partially overlapping
# samples. Journal of Applied Quantitative Methods, 10(3)
The proportions for two groups, "a" and "b" are compared where the raw data "1", or "0" for each unit is recorded in a data frame. The 15 paired observations in this example are given first, followed by 9 independent observations in Sample 1, followed by 6 independent observations in Sample 2.

Independent observations and the paired samples stacked for each sample:

```r
a <- c(1,1,1,1,1,1,1,0,0,0,0,0,1,1,1,1,0,0,0,0,NA,NA,NA,NA,NA,NA,NA)
b <- c(1,1,1,1,1,1,1,0,1,1,0,0,NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,1,1,1,1,1)
```

Prop.test(a, b, stacked=TRUE, conf.level=.95) #Performs the robust 'z8' test. #p.value = 0.053
Index

Partover.test, 2
Prop.test, 3