

Package ‘IncomPair’

January 20, 2025

Type Package

Title Comparison of Means for the Incomplete Paired Data

Version 0.1.0

Description Implements a variety of nonparametric and parametric methods that are commonly used when the data set is a mixture of paired observations and independent samples. The package also calculates and returns values of different tests with their corresponding p-values.

Bhoj, D. S. (1991) <[doi:10.1002/bimj.4710330108](https://doi.org/10.1002/bimj.4710330108)> ``Testing equality of means in the presence of correlation and missing data".

Dubnicka, S. R., Blair, R. C., and Hettmansperger, T. P. (2002) <[doi:10.22237/jmasm/1020254460](https://doi.org/10.22237/jmasm/1020254460)> ``Rank-based procedures for mixed paired and two-sample designs".

Einsporn, R. L. and Habtzghi, D. (2013) <<https://pdfs.semanticscholar.org/89a3/90bafeb2bc41ed4414533cfd5ab84a6b54b6.pdf>> ``Combining paired and two-sample data using a permutation test".

Ekbohm, G. (1976) <[doi:10.1093/biomet/63.2.299](https://doi.org/10.1093/biomet/63.2.299)> ``On comparing means in the paired case with incomplete data on both responses".

Lin, P. E. and Stivers, L. E. (1974) <[doi:10.1093/biomet/61.2.325](https://doi.org/10.1093/biomet/61.2.325)> On difference of means with incomplete data ``.

Maritz, J. S. (1995) <[doi:10.1111/j.1467-842x.1995.tb00649.x](https://doi.org/10.1111/j.1467-842x.1995.tb00649.x)> "A permutation paired test allowing for missing values".

License GPL (>= 2)

Encoding UTF-8

LazyData true

Depends R (>= 2.10)

RoxygenNote 7.0.0

Suggests testthat (>= 2.1.0),

Imports methods

NeedsCompilation no

Author Desale Habtzghi [aut, cre],

Yilin Zhang [aut],

Richard Einsporn [ctb]

Maintainer Desale Habtzghi <dhabtzgh@depaul.edu>

Repository CRAN

Date/Publication 2020-03-05 11:00:02 UTC

Contents

DNA	2
IncomPair	3
parmb	3
permb	4
race	6
rankb	6
Index	8

DNA

DNA Extraction Concentration Data on Coyotes

Description

A total of 30 different coyotes were used for the study. One of the methods was the QIAGEN DNeasy Blood and Tissue Kit and the other was the more traditional chloroform isoamyl alcohol method. Selected randomly, 6 of these coyotes were measured using both methods; 8 were measured with the kit method only and the remaining 16 were measured with chloroform method only.

Usage

DNA

Format

The data frame has 30 rows and 2 variables:

kit DNA measured with the kit method, in ng/uL

chloroform DNA measured with the chloroform method, in ng/uL

Source

Riordan, B. (2012). Northeastern Ohio Coyote Hybridization with Wolves. Honors research project, University of Akron, Akron.

 IncomPair

IncomPair: Comparison of Means for the Incomplete Paired Data

Description

The package contains three functions and two data sets. The functions perform testing the hypothesis of equality of means for the incomplete pairs setting data, including `permb`, `parmb` and `rankb`. `Permb` implements two different nonparametric tests based on a permutation test; `rankb` uses a rank-based procedures; and `parmb` uses different test statistics which are parametric in nature.

Details

For a complete list, use `help(package="IncomPair")`

Author(s)

Desale Habtzghi and Yilin Zhang

Maintainer: Desale Habtzghi <dhabtzgh@depaul.edu>

 parmb

Parametric Based Tests for Incomplete Paired Data

Description

The function performs testing the Hypothesis of equality of means for the incomplete pairs setting data. The function uses six test statistics that were proposed for testing the equality of the means of a bivariate normal distribution with unknown common variance and correlation coefficient when observations are missing on both variates. These function includes Lin and Stivers (1974, Ts), Bhoj (1989, pp. 282, Z), Bhoj (1989, pp. 282, Zb), Bhoj (1989, pp. 283, T), Bhoj (1989, pp. 283, Zh) and Bhoj (1989, pp 284, Zls). For more details, information of the functions see Bhoj (1989).

Usage

```
parmb(xp, yp, xu, yu, r, mu = NULL, method = "Zb",
      alternative = "two.sided", verbose = TRUE)
```

Arguments

<code>xp, yp</code>	(non-empty) numeric vectors of data values of the the complete pairs
<code>xu</code>	a numeric vector of data on x only
<code>yu</code>	a numeric vector of data on y only
<code>r</code>	a number indicating the correlation between the complete pairs
<code>mu</code>	a number indicating the true value of the mean (or difference in means if you are performing a two sample test)

method	a character string specifying the different type of methods, must be one of "Zb" (default), "Zb", "T", "Tls", "Zls", "Zh"
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less"
verbose	if TRUE, show the test statistic used, null and alternative hypotheses in addition to the p-value

Value

A S4 object containing the following components:

Title	a character string describing the test statistic used
Nhypothesis	a character string describing the null hypothesis
Ahypothesis	a character string describing the alternative hypothesis
Pval	the p-value for the test

References

- 1 Bhoj, D. S. (1991). Testing equality of means in the presence of correlation and missing data. *Biometrical journal*, 33(1), 63-72.
- 2 Ekbohm, G. (1976). On comparing means in the paired case with incomplete data on both responses. *Biometrika* 63(2), 299-304.
- 3 Lin, P. E., & Stivers, L. E. (1974). On difference of means with incomplete data. *Biometrika*, 61(2), 325-334.

Examples

```
n=20
n1=15
n2=10
r=0.8
xp=rnorm(n)
yp=r*xp+(1-r)*(rnorm(n))
xu=rnorm(n1)
yu=rnorm(n2)
mu=0
parmb(xp,yp,xu,yu,r,mu,method="Zb",alternative="two.sided")
```

permb

Permutation Based Tests for Incomplete Paired Data

Description

The function Performs testing the hypothesis of equality of means for the incomplete pairs setting data. The function combines the observed mean difference for the complete pairs with the difference between the two means of the independent samples. The function implements two different non-parametric tests based on permutation tests that were proposed by Einsporn and Habtzghi (2013), and Maritz (1995). The two methods are denoted by EH and Maritz, respectively.

Usage

```
permb(xp, yp, xu, yu, r, mu = NULL, method = "EH",
      alternative = "two.sided", verbose = TRUE)
```

Arguments

xp, yp	(non-empty) numeric vectors of data values of the the complete pairs
xu	a numeric vector of data on x only
yu	a numeric vector of data on y only
r	a number indicating the correlation between the complete pairs
mu	a number indicating the true value of the mean (or difference in means if you are performing a two sample test)
method	a character string specifying the different type of methods, must be one of "EH" (default) or "Maritz"
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less"
verbose	if TRUE, show the test used, null and alternative hypotheses in addition to the p-value

Value

A S4 object containing the following components:

Title	a character string describing the test used
Nhypothesis	a character string describing the null hypothesis
Ahypothesis	a character string describing the alternative hypothesis
Pval	the p-value for the test

References

- 1 Einsporn, R. L., & Habtzghi, D. (2013). Combining paired and two-sample data using a permutation test. *Journal of Data Science*, 11(4), 767-779.
- 2 Maritz, J. S. (1995). A permutation paired test allowing for missing values. *Australian Journal of Statistics*, 37(2), 153-159.
- 3 Woolson, R., Leeper, J., Cole, J. and Clarke, W. (1976). A Monte Carlo investigation of a statistic for a bivariate missing data problem. *Communications in Statistics - Theory and Methods A5*, 681-688.

Examples

```
n=20
n1=15
n2=10
r=0.8
xp=rnorm(n)
yp=r*xp+(1-r)*(rnorm(n))
```

```
xu=rnorm(n1)
yu=rnorm(n2)
mu=0
permb(xp,yp,xu,yu,r,mu,method="Maritz",alternative="two.sided")
```

race

5K Race Data

Description

Runner's time for 2 local 5K races in Northeast Ohio (Kent State University 5K race and Tallmadge Memorial 5K) were recorded. There were 32 runners who competed in both of these races, 478 who competed in only the Kent race, and 541 who competed in only the Tallmadge race in 2011.

Usage

```
race
```

Format

The data frame has 1051 observations and 2 variables:

Kent runner's time completing the Kent course, in seconds

Tallmadge runner's time completing the Tallmadge course, in seconds

Source

- 1 <https://ritchessports.com/wp-content/uploads/tallmadge5k/2011-men-overall.txt>
- 2 <https://ritchessports.com/wp-content/uploads/tallmadge5k/2011-women-overall.txt>
- 3 <https://www.kent.edu/RunTheWorld/past-races>

rankb

Rank Based Tests for Incomplete Paired Data

Description

The function performs testing the hypothesis of equality of means for the incomplete pairs setting data. The function uses a rank-based procedure for parameter estimation and hypothesis testing when the data are a mixture of paired observations and independent samples. The rank-based methods combine Wilcoxon signed-rank statistics and Wilcoxon-Mann-Whitney two-sample procedures. These methods were developed by Dubnicka, Blair and Hettmansperger (2002).

Usage

```
rankb(xp, yp, xu, yu, mu = NULL, method = "Ranku",
      alternative = "two.sided", verbose = TRUE)
```

Arguments

xp, yp	(non-empty) numeric vectors of data values of the the complete pairs
xu	a numeric vector of data on x only
yu	a numeric vector of data on y only
mu	a number indicating the true value of the mean (or difference in means if you are performing a two sample test)
method	a character string specifying the different type of methods, must be one of "Ranku" or "Rankw"
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less"
verbose	if TRUE, show the test used, null and alternative hypotheses in addition to the p-value

Value

A S4 object containing the following components:

Title	a character string describing the test used
Nhypothesis	a character string describing the null hypothesis
Ahypothesis	a character string describing the alternative hypothesis
Pval	the p-value for the test

References

Dubnicka, S. R., Blair, R. C., & Hettmansperger, T. P. (2002). Rank-based procedures for mixed paired and two-sample designs. *Journal of Modern Applied Statistical Methods*, 1(1), 6.

Examples

```
n=20
n1=15
n2=10
r=0.8
xp=rnorm(n)
yp=r*xp+(1-r)*(rnorm(n))
xu=rnorm(n1)
yu=rnorm(n2)
mu=0
rankb(xp,yp,xu,yu,mu,method="Rankw",alternative="two.sided")
```

Index

* **datasets**

DNA, [2](#)

race, [6](#)

DNA, [2](#)

IncomPair, [3](#)

parmb, [3](#)

permb, [4](#)

race, [6](#)

rankb, [6](#)