Package ‘clust.bin.pair’

October 26, 2016

Title Statistical Methods for Analyzing Clustered Matched Pair Data
Version 0.0.6
Description
Tests, utilities, and case studies for analyzing significance in clustered binary matched-pair
data. The central function clust.bin.pair uses one of several tests to calculate a Chi-square
statistic. Implemented are the tests Eliasziw, Obuchowski, Durkalski, and Yang with McNemar
included for comparison. The utility functions nested.to.contingency and paired.to.contingency
convert data between various useful formats. Thyroids and psychiatry are the canonical
datasets from Obuchowski and Petryshen respectively.

Depends R (>= 3.2.4)
License MIT + file LICENSE
Encoding UTF-8
LazyData true
URL https://github.com/dgopstein/clust.bin.pair
BugReports https://github.com/dgopstein/clust.bin.pair/issues
Imports
Suggests testthat
RoxygenNote 5.0.1
NeedsCompilation no
Author Dan Gopstein [aut, cre]
Maintainer Dan Gopstein <dan@gopstein.com>
Repository CRAN
Date/Publication 2016-10-26 10:37:47

R topics documented:

clust.bin.pair ................................................................. 2
nested.to.contingency ................................................. 3
obfuscation ................................................................. 4
Statistical test for clustered binary matched pair data

Description

A single interface for several adjustments to the mcnemar test for marginal homogeneity that correct for clustered data.

Usage

clust.bin.pair(ak, bk, ck, dk, method = "yang")

Arguments

ak vector containing counts per group of Success/Success results.
bk vector containing counts per group of Success/Failure results.
ck vector containing counts per group of Failure/Success results.
dk vector containing counts per group of Failure/Failure results.
method a character string specifying the method to calculate the statistic. Must be one of "yang" (default), "durkalski", "obuchowski", "eliasziw". A value of "mcnemar" can also be supplied for comparison.

Value

A list with class "htest" containing the following components:

statistic the value of the test statistic.
p.value the p-value for the test.
method the type of test applied.
data.name a character string giving the names of the data.

References


Examples

tc <- data.frame(nested.to.contingency(thyroids$x.pet, thyroids$x.spect))
clust.bin.pair(tc$ak, tc$bk, tc$ck, tc$dk, method="obuchowski")

pc <- psychiatry[, c('ah', 'bh', 'ch', 'dh')]
clust.bin.pair(pc$ah, pc$bh, pc$ch, pc$dh, method="eliasziw")

nested.to.contingency Convert between nested results and the canonical contingency tables

Description

Sum all concordant and discordant pairs from each nested group into a contingency table.

Usage

nested.to.contingency(t1, t2)

Arguments

t1 lists of pre-treatment measures
t2 lists of post-treatment measures

Value

Contingency tables represented in the rows of a matrix

Examples

nested.to.contingency(thyroids$x.pet, thyroids$x.spect)
**Description**

Data from Gopstein et. al.'s experiment on the misinterpretation of C code. Subjects were asked to hand evaluate pairs of functionally equivalent code. Half of the questions were intentionally obfuscated to elicit confusion.

**Usage**

data(obfuscation)

**Format**

A data frame with 57 rows and 4 variables:

- **subject** the ID of the study participant
- **atom** the type of obfuscation being evaluated
- **control** whether the subject answered the un-obfuscated question correctly
- **treatment** whether the subject answered the obfuscated question correctly

**Source**

Atoms of Confusion

**Examples**

data(obfuscation)

```r
oc <- paired.to.contingency(group = obfuscation[, c("subject", "atom")],
                           t1 = obfuscation$control,
                           t2 = obfuscation$treatment)

clust.bin.pair(oc$sak, oc$sbk, oc$sk, oc$dk, method="durkalski")
```

---

**Description**

Convert between paired results and the canonical contingency tables

**Usage**

```r
paired.to.contingency(group, t1, t2)
```
Arguments

- **group**: List of grouping values
- **t1**: pre-treatment measures
- **t2**: post-treatment measures

Value

Contingency tables represented in the rows of a matrix

Examples

```r
paired.to.contingency(list(obfuscation$subject, obfuscation$atom),
                      obfuscation$control, obfuscation$treatment)
```

Description

Psychiatrists and their patients were surveyed in pairs regarding patient concerns and treatment. Each psychiatrist was asked whether each question item was relevant to their patient and each of their patients were asked the same. The data can be evaluated to answer the question of whether there was patient/doctor agreement in each item. The sample was 29 psychiatrists, each with 1-8 patients, for a total of N = 135 matched pairs.

Usage

```r
data(psychiatry)
```

Format

A data frame with 29 rows and 7 variables:

- **psychiatrist**: the ID of the psychiatrist
- **Nh**: the number of the psychiatrist’s patients participating in the experiment
- **ah**: both participants answered 1
- **bh**: patient answered 1, psychiatrist answered 0
- **ch**: patient answered 0, psychiatrist answered 1
- **dh**: both participants answered 0
- **Wh**: Normalized difference: \( \frac{bh - ch}{Nh} \)
Source


References


Examples

data(psychiatry)

psychiatry$wh == round((psychiatry$bh - psychiatry$ch) / psychiatry$nh, 2)

clust.bin.pair(psyhiatrty$ah, psychiatry$bh, psychiatry$ch, psychiatry$dh, method="eliasziw")

---

thyroids

*PET and SPECT data for diagnosing hyperparathyroidism*

Description

Following surgery which confirmed the absence of hyperparathyroidism two diagnostic tests, PET and SPECT, were performed. Their measures of true negatives and false positives are reported. Data reported in Obuchowski 1998.

Usage

data(thyroids)

Format

A data frame with 21 rows and 6 variables:

- **patient**: ID of the patient
- **n.glands**: number of glands tested from the patient
- **n.pet**: number of true negatives from the PET test
- **x.pet**: individual results per gland from the PET test
- **n.spect**: number of true negatives from the SPECT test
- **x.spect**: individual results per gland from the SPECT test

Source

**Examples**

```r
data(thyroids)

thyroids$n.glands <- sapply(thyroids$x.pet, length)
thyroids$n.glands <- sapply(thyroids$x.spect, length)

thyroids$n.pet <- sapply(thyroids$x.pet, function(x) length(which(x == 1)))
thyroids$n.spect <- sapply(thyroids$x.spect, function(x) length(which(x == 1)))

tc <- nestedNtoNcontingency(thyroids$x.pet, thyroids$x.spect)

clust.bin.pair(tc[, 'ak'], tc[, 'bk'], tc[, 'ck'], tc[, 'dk'], method="obuchowski")
do.call(clust.bin.pair, data.frame(tc))
```
Index

*Topic datasets
  obfuscation, 4
  psychiatry, 5
  thyroids, 6

clust.bin.pair, 2
nested.to.contingency, 3
obfuscation, 4
paired.to.contingency, 4
psychiatry, 5
thyroids, 6