Package 'samplesizeCMH'

December 21, 2017

Title Power and Sample Size Calculation for the
Cochran-Mantel-Haenszel Test

Date 2017-12-13

Version 0.0.0

Copyright Spectrum Health, Grand Rapids, MI

Description Calculates the power and sample size for Cochran-Mantel-Haenszel tests.
There are also several helper functions for working with probability,
odds, relative risk, and odds ratio values.

Depends R (>= 3.4.0)

Imports stats

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

Suggests knitr, rmarkdown, DescTools, datasets, testthat

VignetteBuilder knitr

RoxygenNote 6.0.1

URL https://github.com/pegeler/samplesizeCMH

BugReports https://github.com/pegeler/samplesizeCMH/issues

NeedsCompilation no

Author Paul Egeler [aut, cre],
Spectrum Health, Grand Rapids, MI [cph]

Maintainer Paul Egeler <paul.egeler@spectrumhealth.org>

Repository CRAN

Date/Publication 2017-12-21 15:15:57 UTC

R topics documented:

---

contraceptives	*Oral contraceptive use and breast cancer rates*

Description

This data summarizes counts of a case-control study investigating the link between breast cancer rates and oral contraceptive use, stratified by age group. In toto, 10,890 subjects. See source for details.

Usage

```
data(contraceptives)
```

Format

A 3-dimensional table.

1. `OC Usage`: Subject exposure to oral contraceptives.

2. `Disease Status`: Breast cancer present (case) or absent (control).

3. `Age Group`: Age group of the subject.

Source

Hennekens, C. H., F. E. Speizer, R. J. Lipnick, B. Rosner, C. Bain, C. Belanger, M. J. Stampfer, W. Willett, and R. Peto. (1984). "A Case-Control Study of Oral Contraceptive Use and Breast Cancer." *Journal of the National Cancer Institute* **72** (1): 39–42. Table 1.

odds.and.proportions   *Interconvert odds and proportion values*

Description

These functions will create either odds for a given probability, probability for a given odds, calculate the odds ratio between two probabilities, or calculate effect size (raise a probability by theta)

Usage

```
prop2odds(p)

odds2prop(o)

effect.size(p, theta)

props2theta(p1, p2)

rr2theta(rr, p1, p2)

theta2rr(theta, p1, p2)
```

Arguments

p, p1, p2	Proportion vector.
o	Odds vector.
theta	Odds ratio vector.
rr	Relative risk vector (p1 / p2).

Value

A numeric vector.

Author(s)

Paul W. Egeler, M.S.

Examples

```
# Convert proportions of 0 through 1 to odds
props <- seq(0,1,0.1)
prop2odds(props)

# Convert odds to proportions
odds2prop(1:3)

# Raise a proportion by an effect size theta
```

```
effect.size(0.5, 2)

# Find the odds ratio between to proportions
props2theta(0.75, 0.5)
```

---

odds.ratio	*Create an odds ratio estimate from a 2-by-2 table of frequencies or proportions*

Description

Create an odds ratio estimate from a 2-by-2 table of frequencies or proportions

Usage

```
odds.ratio(x)
```

Arguments

x   A two-dimensional matrix or table containing frequencies or proportions.

Value

A numeric vector.

Author(s)

Paul W. Egeler, M.S.

Examples

```
# Load in Titanic data from datasets package
data(Titanic, package = "datasets")

# Get marginal table of survival by sex
marginal_table <- margin.table(Titanic, c(2,4))
marginal_table

# Compute odds ratio of marginal table
odds.ratio(marginal_table)

# Get partial tables of survival by sex, stratified by class
partial_tables <- margin.table(Titanic, c(2,4,1))
partial_tables

# Compute odds ratio of each partial table
apply(partial_tables, 3, odds.ratio)
```

---

power.cmh.test *Power and sample size calculation for the Cochran-Mantel-Haenszel test*

Description

Compute the post-hoc power or required number of subjects for the Cochran-Mantel-Haenszel test for association in *J* stratified 2 x 2 tables.

Usage

```
power.cmh.test(p1 = NULL, p2 = NULL, theta = NULL, N = NULL,
  sig.level = 0.05, power = 0.8, alternative = c("two.sided", "less",
  "greater"), s = 0.5, t = 1/J, correct = TRUE)
```

Arguments

p1	Vector of proportions of the *J* case groups.
p2	Vector of proportions of the *J* control groups.
theta	Vector of odds ratios relating to the *J* 2 x 2 tables.
N	Total number of subjects.
sig.level	Significance level (Type I error probability).
power	Power of test (1 minus Type II error probability).
alternative	Two- or one-sided test. If one-sided, the direction of the association must be defined (less than 1 or greater than 1). Can be abbreviated.
s	Proportion (weight) of case versus control in *J* stratum.
t	Proportion (weight) of total number of cases of *J* stratum.
correct	Logical indicating whether to apply continuity correction.

Details

This sample size calculation is based on the derivations described in the Woolson *et al.* (1986). It is designed for case-control studies where one margin is fixed. The method is "based on the Cochran-Mantel-Haenszel statistic expressed as a weighted difference in binomial proportions."

Continuity corrected sample size is described in Nam's 1992 paper. This uses the weighted binomial sample size calculation described in Woolson *et al.* (1986) but is enhanced for use with the continuity corrected Cochran's test.

Power calculations are based on the writings of Wittes and Wallenstein (1987). They are functionally equivalent to the derivations of the sample size calculation described by Woolson and others and Nam, but have slightly added precision.

Terminology and symbolic conventions are borrowed from Woolson *et al.* (1986). The p1 group is dubbed the *Case* group and p2 group is called the *Control* group.

Value

An object of class `"power.cmh"`: a list of the original arguments and the calculated sample size or power. Also included are vectors of n's per each group, an indicator or whether continuity correction was used, the original function call, and `N.effective`.

The vectors of n's per each group, `n1` and `n2`, are the fractional n's required to achieve a final total N specified by the calculation while satisfying the constraints of `s` and `t`. However, the effective N, given the requirement of cell counts populated by whole numbers is provided by `N.effective`. By default, the print method is set to `n.frac = FALSE`, which will round each cell n up to the nearest whole number.

Arguments

To calculate **power**, the power parameter must be set to `NULL`. To calculate **sample size**, the `N` parameter must be set to `NULL`.

The J number of groups will be inferred by the maximum length of `p1`, `p2`, or `theta`.

Effect size must be specified using one of the following combinations of arguments.

- Both case and control proportion vectors, ex.,

    - `p1` and `p2` with `theta = NULL`.

- One proportion vector and an effect size, ex.,

    - `p1` and `theta` with `p2 = NULL`, **or**
    - `p2` and `theta` with `p1 = NULL`.

Author(s)

Paul W. Egeler, M.S.

References

Gail, M. (1973). "The determination of sample sizes for trials involving several 2 x 2 tables." *Journal of Chronic Disease* **26**: 669-673.

Munoz, A. and B. Rosner. (1984). "Power and sample size for a collection of 2 x 2 tables." *Biometrics* **40**: 995-1004.

Nam, J. (1992). "Sample size determination for case-control studies and the comparison of stratified and unstratified analyses." *Biometrics* **48**: 389-395.

Wittes, J. and S. Wallenstein. (1987). "The power of the Mantel-Haenszel test." *Journal of the American Statistical Association* **82**: 1104-1109.

Woolson, R. F., Bean, J. A., and P. B. Rojas. (1986). "Sample size for case-control studies using Cochran's statistic." *Biometrics* **42**: 927-932.

See Also

power.prop.test, mantelhaen.test, BreslowDayTest

Examples

```
# From "Sample size determination for case-control studies and the comparison
# of stratified and unstratified analyses", (Nam 1992). See references.

# Uncorrected sample size estimate first introduced
# by Woolson and others in 1986
sample_size_uncorrected <- power.cmh.test(
  p2 = c(0.75,0.70,0.65,0.60),
  theta = 3,
  power = 0.9,
  t = c(0.10,0.40,0.35,0.15),
  alternative = "greater",
  correct = FALSE
)

print(sample_size_uncorrected, detail = FALSE)

# We see that the N is 171, the same as calculated by Nam
sample_size_uncorrected$N


# Continuity corrected sample size estimate added by Nam
sample_size_corrected <- power.cmh.test(
  p2 = c(0.75,0.70,0.65,0.60),
  theta = 3,
  power = 0.9,
  t = c(0.10,0.40,0.35,0.15),
  alternative = "greater",
  correct = TRUE
)

print(sample_size_corrected, n.frac = TRUE)

# We see that the N is indeed equal to that which is reported in the paper
sample_size_corrected$N
```

---

print.power.cmh	*Print power.cmh object*

Description

Print power.cmh object

Usage

```
## S3 method for class 'power.cmh'
print(x, detail = TRUE, n.frac = FALSE, ...)
```

Arguments

x	A ″power.cmh″ object.
detail	Logical to toggle detailed or simple output.
n.frac	Logical indicating whether sample n's should be rounded to the next whole number.
...	Ignored.

---

rel.risk   *Calculate the relative risk from a 2-by-2 table*

Description

Computes the relative risk of a specified column of a two-by-two table.

Usage

```
rel.risk(x, col.num = 1)
```

Arguments

x	A table or matrix containing frequencies.
col.num	The column number upon which relative risk should be calculated.

Value

A numeric vector.

Author(s)

Paul W. Egeler, M.S.

---

samplesizeCMH   *samplesizeCMH: Power and Sample Size Calculation for the Cochran-Mantel-Haenszel Test*

Description

This package provides functions relating to power and sample size calculation for the CMH test. There are also several helper functions for interconverting probability, odds, relative risk, and odds ratio values.

Details

The **Cochran-Mantel-Haenszel test** (CMH) is an inferential test for the association between two binary variables, while controlling for a third confounding nominal variable. Two variables of interest, *X* and *Y*, are compared at each level of the confounder variable *Z* and the results are combined, creating a common odds ratio. Essentially, the CMH test examines the *weighted* association of *X* and *Y*. The CMH test is a common technique in the field of biostatistics, where it is often used for case-control studies.

Sample Size Calculation

Given a target power which the researcher would like to achieve, a calculation can be performed in order to estimate the appropriate number of subjects for a study. The `power.cmh.test` function calculates the required number of subjects per group to achieve a specified power for a Cochran-Mantel-Haenszel test.

Power Calculation

Researchers interested in estimating the probability of detecting a true positive result from an inferential test must perform a power calculation using a known sample size, effect size, significance level, *et cetera*. The `power.cmh.test` function can compute the power of a CMH test, given parameters from the experiment.

Index