Package 'Exact'

October 14, 2019

Type Package

Title Unconditional Exact Test

Version 2.0

Author Peter Calhoun [aut, cre]

Maintainer Peter Calhoun <calhoun.peter@gmail.com>

Description Performs unconditional exact tests and power calculations for 2x2 contingency tables.

License GPL-2

Depends R (>= 3.1.0)

Imports graphics, stats

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2019-10-14 04:40:02 UTC

R topics documented:

Exact-package   *Unconditional Exact Tests R Package*

Description

This package performs unconditional exact tests using the `exact.test` function. This package also includes the `power.exact.test` function to calculate the power of various tests and the `exact.reject.region` to determine the rejection region for a given sample size.

Details

Unconditional exact tests are a more powerful alternative than conditional exact tests. This package can compute p-values, confidence intervals, and power calculations for various tests. Details of the tests are given in the exact.test documentation.

Note

Thoughout the years I have received help while creating this package. Special thanks goes to Philo Calhoun, Tal Galili, Kamil Erguler, Roger Berger, Karl Hufthammer, and the R community.

Author(s)

Peter Calhoun [aut, cre]

Maintainer: Peter Calhoun <calhoun.peter@gmail.com>

---

exact.reject.region   *Rejection Region for 2x2 Tables*

Description

Determines the rejection region for known sample sizes and significance level.

Usage

```
exact.reject.region(n1, n2, alternative = c("two.sided", "less", "greater"),
    alpha = 0.05, npNumbers = 100, np.interval = FALSE, beta = 0.001,
    method = c("z-pooled", "z-unpooled", "boschloo", "santner and snell",
                "csm", "csm approximate", "fisher", "chisq", "yates chisq"),
    ref.pvalue = TRUE, delta = 0, convexity = TRUE)
```

Arguments

n1	The sample size in first group
n2	The sample size in second group
alternative	Indicates the alternative hypothesis: must be either "two.sided", "less", or "greater"
alpha	Significance level
npNumbers	Number: The number of nuisance parameters considered
np.interval	Logical: Indicates if a confidence interval on the nuisance parameter should be computed
beta	Number: Confidence level for constructing the interval of nuisance parameters considered. Only used if np.interval=TRUE
method	Indicates the method for finding the more extreme tables: must be either "Z-pooled", "Z-unpooled", "Santner and Snell", "Boschloo", "CSM", "CSM approximate", "Fisher", "Chisq", or "Yates Chisq"

ref.pvalue	Logical: Indicates if p-value should be refined by maximizing the p-value function after the nuisance parameter is selected
delta	Number: null hypothesis of the difference in proportion. Only used if method is "z-pooled" or "csm"
convexity	Logical: assumes convexity for interval approach. Only used if np.interval=TRUE

Details

The rejection region are calculated for binomial models. The design must know the fixed sample sizes in advance. Rejection region can be determined for any `exact.test` computation, Fisher's exact test, or chi-square test (Yates' or Pearson's). All tests attain the convexity property which greatly speeds up computation time (see Chan for definition). However, in very rare cases, using the nuisance parameter interval approach does not attain the convexity property, so it is possible using `convexity=TRUE` could yield an inaccurate power calculation with this method. This is *extremely unlikely* though, so default is to assume convexity and speed up computation time.

Value

A matrix of the rejection region. The columns represent the number of successes in first group, rows represent the number of successess in second group, and cells represent whether the test is rejected (1) or failed to be rejected (0). This matrix represents all possible 2x2 tables.

Note

Not refining the p-value often yields similar results and decreases the computation time.

Author(s)

Peter Calhoun

References

Barnard, G.A. (1947) Significance tests for 2x2 tables. *Biometrika*, **34**, 123-138

Chan, I. (2003), Proving non-inferiority or equivalence of two treatments with dichotomous end-points using exact methods, *Statistical Methods in Medical Research*, **12**, 37–58

See Also

power.exact.test

Examples

```
exact.reject.region(n1=10, n2=20, alternative="two.sided", method="Z-pooled")
exact.reject.region(n1=10, n2=20, alternative="less", method="CSM", delta=0.10)
```

| exact.test | *Unconditional Exact Tests for 2x2 tables* |

Description

Calculates Barnard's or Boschloo's unconditional exact test for binomial or multinomial models

Usage

```
exact.test(data, alternative = c("two.sided", "less", "greater"), npNumbers = 100,
    np.interval = FALSE, beta = 0.001,
    method = c("z-pooled", "z-unpooled", "boschloo", "santner and snell",
               "csm", "csm approximate"),
    model = c("Binomial", "Multinomial"), conf.int = FALSE, conf.level = 0.95,
    precision = 0.001, cond.row = TRUE, to.plot = TRUE, ref.pvalue = TRUE,
    delta = 0, reject.alpha = NULL)
```

Arguments

data	A two dimensional contingency table in matrix form
alternative	Indicates the alternative hypothesis: must be either "two.sided", "less", or "greater"
npNumbers	Number: The number of nuisance parameters considered
np.interval	Logical: Indicates if a confidence interval on the nuisance parameter should be computed
beta	Number: Confidence level for constructing the interval of nuisance parameters considered. Only used if np.interval=TRUE
method	Indicates the method for finding the more extreme tables: must be either "Z-pooled", "Z-unpooled", "Santner and Snell", "Boschloo", "CSM", or "CSM approximate". CSM tests cannot be calculated for multinomial models
model	The model being used: must be either "Binomial" or "Multinomial"
conf.int	Logical: Indicates if a confidence interval on the difference in proportion should be computed. Only used if model="Binomial" and method is "z-pooled" or "csm"
conf.level	Number: Confidence level of interval on difference in proportion. Only used if conf.int=TRUE
precision	Number: Precision of confidence interval on difference in proportion. Smaller number represents more digits in confidence interval. Only used if conf.int=TRUE
cond.row	Logical: Indicates if row margins are fixed in the binomial models. Only used if model="Binomial"
to.plot	Logical: Indicates if plot of p-value vs. nuisance parameter should be generated. Only used if model="Binomial"
ref.pvalue	Logical: Indicates if p-value should be refined by maximizing the p-value function after the nuisance parameter is selected. Only used if model="Binomial"

delta   Number: null hypothesis of the difference in proportion. Only used if model="Binomial" and method is "z-pooled" or "csm"

reject.alpha   Number: significance level for exact test. Optional and primarily used to speed up exact.reject.region function

Details

Unconditional exact tests can be performed for binomial or multinomial models. The binomial model assumes the row or column margins (but not both) are known in advance, while the multinomial model assumes only the total sample size is known beforehand. For the binomial model, the user needs to specify which margin is fixed (default is rows). Conditional tests (e.g., Fisher's exact test) have both row and column margins fixed, but this is a very uncommon design.

For the binomial model, the null hypothesis is the difference of proportion is equal to 0. Under the null hypothesis, the probability of a 2x2 table is the product of two binomials. The p-value is calculated by maximizing a nuisance parameter and summing the as or more extreme tables. The method parameter specifies the method to determine the more extreme tables (see references for more details):

- Z-pooled (or Score) - Uses the test statistic from a Z-test using a pooled proportion

- Z-unpooled (or Wald) - Uses the test statistic from a Z-test without using the pooled proportion

- Santner and Snell - Uses the difference in proportion

- Boschloo - Uses the p-value from Fisher's exact test

- CSM - Starts with the most extreme table and sequentially adds more extreme tables based on the smallest p-value (calculated by maximizing the probability of a 2x2 table). This is Barnard's original method

- CSM approximate - Similar to the CSM test, but does not update the maximum p-value calculation when each table is added

There is some disagreement on which method to use. Suissa and Shuster suggested using a Z-pooled statistic, which is uniformly more powerful than Fisher's test for balanced designs. Boschloo recommended using the p-value for Fisher's test as the test statistic. This method became known as Boschloo's test, and it is always uniformly more powerful than Fisher's test. Mato and Andres suggested using Barnard's CSM test. This is the recommended method by the author of this R package, but is much more computationally intensive. Regardless of the method chosen, most researchers agree that Fisher's exact test should **not** be used to analyze 2x2 tables.

Once the more extreme tables are determined, the p-value is calculated the maximizing over the common success probability – a nuisance parameter. The p-value computation has many local maxima and can be computationally intensive. The code performs an exhaustive search by considering many values of the nuisance parameter from 0 to 1, represented by npNumbers. Another approach, proposed by Berger and Boos, is to calculate the Clopper-Pearson confidence interval of the nuisance parameter (represented by np.interval) and only maximizing the p-value function for nuisance parameters within the confidence interval; this approach adds a small penalty to the p-value to control for the type 1 error rate (cannot be used with CSM). If ref.pvalue = TRUE, then the code will also use the optimise function near the nuisance parameter to refine the p-value. Increasing npNumbers and using ref.pvalue ensures the p-value is correctly calculated at the expense of slightly more computation time.

The Z-pooled and CSM test has been extended to test the difference in proportion is equal to `delta` (not only 0). Thus, the Z-pooled and CSM test can be used to test noninferiority designs, and be used to construct confidence intervals for the difference in proportion.

There are many ways to define the two-sided p-value; this code uses the `fisher.test` approach by summing the probabilities for both sides of the table. Confidence intervals of the difference in proportion (Z-pooled and CSM test only) use the Agresti-Min interval approach for two-sided tests and the Chan-Zhang interval approach for one-sided tests. The `precision` represents the number of decimal points for the confidence interval.

The above description applies to the binomial model. The multinomial model is similar except there are two nuisance parameters. The CSM test has not been developed for multinomial models nor have noninferiority tests been developed. Improvements to the code have focused on the binomial model, so multinomial models takes substantially longer.

Value

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

`p.value`	The computed p-value
`test.statistic`	The observed test statistic
`estimate`	An estimate of the parameter tested
`alternative`	A character string describing the alternative hypothesis
`model`	A character string describing the model design ("Binomial" or "Multinomial")
`method`	A character string describing the method to determine more extreme tables
`np`	The nuisance parameter that maximizes the p-value. For multinomial models, both nuisance parameters are given
`np.range`	The range of nuisance parameters considered. For multinomial models, both nuisance parameter ranges are given
`data.name`	A character string giving the names of the data

Warning

Multinomial models and CSM tests may take a very long time, even for small sample sizes.

Note

CSM test and multinomial models are much more computationally intensive. I have also spent a greater amount of time making the computations for the binomial models more efficient; future work will be devoted to improving the multinomial models. Increasing the number of nuisance parameters considered and refining the p-value will increase the computation time. Performing confidence intervals also greatly increases computation time.

This code was influenced by the FORTRAN program located at http://www4.stat.ncsu.edu/~boos/exact/

Author(s)

Peter Calhoun

References

Agresti, A. and Min, Y. (2001) On small-sample confidence intervals for parameters in discrete distributions. *Biometrics*, **57**, 963–971

Barnard, G.A. (1945) A new test for 2x2 tables. *Nature*, **156**, 177

Barnard, G.A. (1947) Significance tests for 2x2 tables. *Biometrika*, **34**, 123–138

Berger, R. and Boos D. (1994) P values maximized over a confidence set for the nuisance parameter. *Journal of the American Statistical Association*, **89**, 1012–1016

Berger, R. (1994) Power comparison of exact unconditional tests for comparing two binomial proportions. *Institute of Statistics Mimeo Series No. 2266*

Berger, R. (1996) More powerful tests from confidence interval p values. *American Statistician*, **50**, 314–318

Boschloo, R. D. (1970), Raised Conditional Level of Significance for the 2x2-table when Testing the Equality of Two Probabilities. *Statistica Neerlandica*, **24**, 1–35

Chan, I. (2003), Proving non-inferiority or equivalence of two treatments with dichotomous endpoints using exact methods, *Statistical Methods in Medical Research*, **12**, 37–58

Cardillo, G. (2009) MyBarnard: a very compact routine for Barnard's exact test on 2x2 matrix. http://www.mathworks.com/matlabcentral/fileexchange/25760-mybarnard

Mato, S. and Andres, M. (1997), Simplifying the calculation of the P-value for Barnard's test and its derivatives. *Statistics and Computing*, **7**, 137–143

Mehrotra, D., Chan, I., Berger, R. (2003), A Cautionary Note on Exact Unconditional Inference for a Difference Between Two Independent Binomial Proportions. *Biometrics*, **59**, 441–450

Ruxton, G. D. and Neuhauser M (2010), Good practice in testing for an association in contingency tables. *Behavioral Ecology and Sociobiology*, **64**, 1505–1513

Suissa, S. and Shuster, J. J. (1985), Exact Unconditional Sample Sizes for the 2x2 Binomial Trial, *Journal of the Royal Statistical Society, Ser. A*, **148**, 317–327

See Also

`fisher.test` and **exact2x2**

Examples

```
data <- matrix(c(7, 8, 12, 3), 2, 2, byrow=TRUE)
exact.test(data, alternative="less",to.plot=TRUE)
exact.test(data, method="Z-pooled", alternative="two.sided", conf.int=TRUE, conf.level=0.95,
           precision=0.0001, npNumbers=100)
exact.test(data, method="Boschloo", alternative="two.sided", np.interval=TRUE, beta=0.001,
           npNumbers=100, to.plot=FALSE)

# Example from Barnard's (1947) appendix:
data <- matrix(c(4, 0, 3, 7), 2, 2,
               dimnames=list(c("Box 1","Box 2"), c("Defective","Not Defective")))
exact.test(data, method="CSM", alternative="two.sided", conf.int=TRUE)

data <- matrix(c(6, 8, 4, 3), 2, 2, byrow=TRUE)
exact.test(data, model="Multinomial", alternative="less", method="Z-pooled")
```

---

power.exact.test	*Power Calculations for 2x2 Tables*

Description

Calculates the power of the design for known sample sizes and true probabilities.

Usage

```
power.exact.test(p1, p2, n1, n2, alternative = c("two.sided", "less", "greater"),
    alpha = 0.05, npNumbers = 100, np.interval = FALSE, beta = 0.001,
    method = c("z-pooled", "z-unpooled", "boschloo", "santner and snell",
               "csm", "csm approximate", "fisher", "chisq", "yates chisq"),
    ref.pvalue = TRUE, simulation = FALSE, nsim = 100, delta = 0,
    convexity = TRUE)
```

Arguments

p1	The probability of success given in first group
p2	The probability of success given in second group
n1	The sample size in first group
n2	The sample size in second group
alternative	Indicates the alternative hypothesis: must be either "two.sided", "less", or "greater"
alpha	Significance level
npNumbers	Number: The number of nuisance parameters considered
np.interval	Logical: Indicates if a confidence interval on the nuisance parameter should be computed
beta	Number: Confidence level for constructing the interval of nuisance parameters considered. Only used if np.interval=TRUE
method	Indicates the method for finding more extreme tables: must be either "Z-pooled", "Z-unpooled", "Santner and Snell", "Boschloo", "CSM", "CSM approximate", "Fisher", "Chisq", or "Yates Chisq"
ref.pvalue	Logical: Indicates if p-value should be refined by maximizing the p-value function after the nuisance parameter is selected
simulation	Logical: Indicates if the power calculation is exact or estimated by simulation
nsim	Number of simulations run. Only used if simulation=TRUE
delta	Number: null hypothesis of the difference in proportion. Only used if method is "z-pooled" or "csm"
convexity	Logical: assumes convexity for interval approach. Only used if np.interval=TRUE

Details

The power calculations are for binomial models. The design must know the fixed sample sizes in advance. There are (n1+1) x (n2+1) possible tables that could be produced. There are two ways to calculate the power: simulate the tables under two independent binomial distributions or determine the rejection region for all possible tables and calculate the exact power. The calculations can be done using any `exact.test` computation, Fisher's exact test, or chi-square tests (Yates' or Pearson's). The power calculations utilize the convexity property, which greatly speeds up computation time (see `exact.reject.region` documentation).

Value

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

n1, n2	The respective sample sizes
p1, p2	The respective proportions
alpha	Significance level
power	Power of the test
alternative	Alternative hypothesis
delta	Null hypothesis of the difference in proportion
method	Method/test

Note

Not refining the p-value often yields similar results and decreases the computation time.

Author(s)

Peter Calhoun

References

Berger, R. (1994) Power comparison of exact unconditional tests for comparing two binomial proportions. *Institute of Statistics Mimeo Series No. 2266*

Berger, R. (1996) More powerful tests from confidence interval p values. *American Statistician*, **50**, 314-318

Boschloo, R. D. (1970), Raised Conditional Level of Significance for the 2x2-table when Testing the Equality of Two Probabilities. *Statistica Neerlandica*, **24**, 1-35

See Also

`exact.reject.region` and **statmod**

Examples

```
power.exact.test(0.20, 0.80, 10, 20, method="Z-pooled", delta=0.1)
power.exact.test(0.20, 0.80, 10, 20, method="Boschloo", np.interval=TRUE)
power.exact.test(0.20, 0.80, 10, 20, method="Fisher")
```

Index