Package 'papci'

September 14, 2021

Type Package

Title Prevalence Adjusted PPV Confidence Interval

Version 0.1.0

Date 2021-09-13

Maintainer Cui Guo <cuguo@foundationmedicine.com>

Description Positive predictive value (PPV) defined as the conditional probability of clinical trial assay (CTA) being positive given Companion diagnostic device (CDx) being positive is a key performance parameter for evaluating the clinical validity utility of a companion diagnostic test in clinical bridging studies. When bridging study patients are enrolled based on CTA assay results, Binomial-based confidence intervals (CI) may are not appropriate for PPV CI estimation. Bootstrap CIs which are not restricted by the Binomial assumption may be used for PPV CI estimation only when PPV is not 100%. Bootstrap CI is not valid when PPV is 100% and becomes a single value of [1, 1]. We proposed a risk ratio-based method for constructing CI for PPV. By simulation we illustrated that the coverage probability of the proposed CI is close to the nominal value even when PPV is high and negative percent agreement (NPA) is close to 100%. There is a lack of R package for PPV CI calculation. we developed a publicly available R package along with this shiny app to implement the proposed approach and some other existing methods.

Imports dplyr, tidyverse, binom, PropCIs, ratesci, Hmisc, shiny, shinythemes, readxl, DT, stats

License GPL-3

RoxygenNote 7.1.1

NeedsCompilation no

Author Cui Guo [aut, cre, ctb],
Kate (Yueyi) Li [aut, ctb],
Lei Yang [aut, ctb],
Chang Xu [aut, ctb]

Repository CRAN

Date/Publication 2021-09-14 07:00:02 UTC

R topics documented:

---

agreement   *agreement*

Description

This function is used to evaluate the agreements of two clinical tests with binary outcomes. It provides four statistical metrics, including positive percent agreement (PPA), negative percent agreement (NPA), positive predictive value (PPV), negative predictive value (NPV) and their corresponding confidence intervals (CIs) in different methods. Some methods available in this package provide non-binomial-based CIs, which are more appropriate for CIs estimations of PPV and NPV when samples are not randomly selected from IU population or PPV=100%.

Usage

```
agreement(x, y, m, n, tb, baseline, comparator, prev, conf.level,
alternative, methods_pa, methods_pv, times, ...)
```

Arguments

x	Number of positive calls by both baseline and comparator tests. x is a non-negative integer.
y	Number of observations called negative at baseline but positive by comparator test. y is a non-negative integer.
m	Number of positive calls at baseline. m is a non-negative integer.
n	Number of negative calls at baseline. n is a non-negative integer.
tb	A 2-by-2 contingency table between baseline and comparator testing results. Baseline at column.
baseline	A vector of baseline testing results. 1 for call, 0 for non-call. The two vectors, baseline and comparator, are paired non-missing results and their length should be the same; ignored otherwise.
comparator	A vector of comparator testing results. 1 for call, 0 for non-call. Same requirements as for baseline.
prev	Disease prevalence based on baseline test in IU population. prev=Pr(baseline negative)
conf.level	The level of confidence to be used in the confidence interval.
alternative	a character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less".

methods_pa	Which method to use to construct the CIs for PPA and NPA. Any combination of `c("exact","ac","asymptotic","wilson","prop.test","bayes","logit","cloglog","probit")` is allowed. Default is `"all"`.
methods_pv	Which method to use to construct the interval for PPV and NPV. Any combination of `c("Koopman","Katz","Neother","Gart_Nam","Bootstrap","Plug-In")` is allowed. Default is `"all"`.
times	Times of bootstraps if `"Bootstrap"` method is specified to construct the interval for PPV. Default is `"1000"`.
...	Additional arguments to be passed to `binom.confint`.

Details

The definitions of PPA, NPA, PPV and NPA in this package are

- $PPA = Pr(comparator + |baseline+)$, $NPA = Pr(comparator - |baseline-)$
- $PPV = Pr(baseline + |comparator+)$, $NPV = Pr(baseline - |comparator-)$

The point estimations are x/m for PPA and and $(n-y)/n$ for NPA. By ignoring enrollment biases, PPV and NPV are estimated as $x/(x+y)$ and $(n-y)/(m+n-x-y)$, respectively. When samples are not enrolled randomly or selected based on baseline results, PPV and NPV are obtained by the Bayes theorem and not binomially distributed. They are defined as

$$PPV = prev * PPA/[prev * PPA + (1 - prev) * (1 - NPA)]$$

$$NPV = (1 - prev) * NPA/[(1 - prev) * NPA + prev * (1 - PPA)]$$

Nine methods are allowed for constructing the confidence interval(s) for PPA and NPA referring to `binom.confint`. Six methods are allowed for constructing the confidence interval(s) for PPV based on the risk-ratio $R1 = (1 - NPA)/PPA$.

- Koopman (1984) - derived the $(1 - \alpha)100\%$ CI for $R1$ by using Chi-squared method.
- Katz et al.(1978) - derived the $(1 - \alpha)100\%$ CI for $R1$ by assuming that the $log(R1)$ is approximately normally distributed.
- Noether (1957) - developed the $(1 - \alpha)100\%$ CI for $R1$ using an explicit solution.
- Gart and Nam (1988) - improved Koopman's method by correcting the asymptotic skewness.
- Bootstrap - derived the risk ratio CI using Bootstrap method from multiple random samples.
- Plug-In - derived the 95% CI for PPA and NPA as (PPA_l,PPA_u) and (NPA_l,NPA_u). Applied all four combinations (i.e., (PPA_l, NPA_l); (PPA_l, NPA_u); (PPA_u, NPA_l); (PPA_u, NPA_u)) into above PPV formulas by Bayes theorem, and the minimum and maximum values are determined as the lower and upper bound of 95% CIs of PPV.

Given the CIs for the risk-ratio $R1$, denoted as $[R1_l, R1_u]$, the CIs for PPV can be directly contained by

$$[p/(p + (1 - p) * R1_u), p/(p + (1 - p) * R1_l)]$$

CIs of NPV can be derived in the same way.

Value

A `list` of `data.frame` containing the estimated agreements (ppa, npa, ppv, npv) and the lower and upper bounds of the confidence interval for all the methods in `methods_pa` or `methods_pv`.

Author(s)

Lei Yang, Cui Guo, Kate Li, Chang Xu (cuguo@foundationmedicine.com)

References

1. Gart John J and Nam Jun-mo (1988). Approximate interval estimation of the ratio of binomial parameters: a review and corrections for skewness, *Biometrics*, 323-338.

2. Katz DJSM, Baptista J, Azen SP and Pike MC (1978). Obtaining confidence intervals for the risk ratio in cohort studies, *Biometrics*, 469-474.

3. Koopman PAR (1984). Confidence intervals for the ratio of two binomial proportions, *Biometrics*, 513-517.

4. Noether Gottfried E (1957). Two confidence intervals for the ratio of two probabilities and some measures of effectiveness, *Journal of the American Statistical Association*, **52**: 36-45.

See Also

[binom.confint](#) for different methods to obtain a confidence interval on the binomial probability like PPA and NPA.

Examples

```
agreement(x = 90, y = 10, m = 100, n = 80, prev = 0.3, times=1000)
agreement(x = 84, y = 0, m = 84, n = 97, prev = 0.096, times=1000)
```

---

papciUI   *papciUI*

Description

This function will automatically launch the PA interactive user interface in a web browser. The user interface can also be accessed by https://kate-yueyi-li.shinyapps.io/shiny . Neither R nor any packages are required in this online version.

Usage

```
papciUI()
```

Details

The definitions of PPA, NPA, PPV and NPA in this package are

- $PPA = Pr(comparator + |baseline+)$, $NPA = Pr(comparator - |baseline-)$
- $PPV = Pr(baseline + |comparator+)$, $NPV = Pr(baseline - |comparator-)$

The point estimations are x/m for PPA and and $(n-y)/n$ for NPA. By ignoring enrollment biases, PPV and NPV are estimated as $x/(x+y)$ and $(n-y)/(m+n-x-y)$, respectively. When samples are not enrolled randomly or selected based on baseline results, PPV and NPV are obtained by the Bayes theorem and not binomially distributed. They are defined as

$$PPV = prev * PPA/[prev * PPA + (1 - prev) * (1 - NPA)]$$

$$NPV = (1 - prev) * NPA/[(1 - prev) * NPA + prev * (1 - PPA)]$$

Nine methods are allowed for constructing the confidence interval(s) for PPA and NPA referring to `binom.confint`. Six methods are allowed for constructing the confidence interval(s) for PPV based on the risk-ratio $R1 = (1 - NPA)/PPA$.

- Koopman (1984) - derived the $(1 - \alpha)100\%$ CI for $R1$ by using Chi-squared method.
- Katz et al.(1978) - derived the $(1 - \alpha)100\%$ CI for $R1$ by assuming that the $log(R1)$ is approximately normally distributed.
- Noether (1957) - developed the $(1 - \alpha)100\%$ CI for $R1$ using an explicit solution.
- Gart and Nam (1988) - improved Koopman's method by correcting the asymptotic skewness.
- Bootstrap - derived the risk ratio CI using Bootstrap method from multiple random samples.
- Plug-In - derived the 95% CI for PPA and NPA as (PPA_l,PPA_u) and (NPA_l,NPA_u). Applied all four combinations (i.e., (PPA_l, NPA_l); (PPA_l, NPA_u); (PPA_u, NPA_l); (PPA_u, NPA_u)) into above PPV formulas by Bayes theorem, and the minimum and maximum values are determined as the lower and upper bound of 95% CIs of PPV.

Given the CIs for the risk-ratio $R1$, denoted as $[R1_l, R1_u]$, the CIs for PPV can be directly contained by

$$[p/(p + (1 - p) * R1_u), p/(p + (1 - p) * R1_l)]$$

CIs of NPV can be derived in the same way.

Value

A `list` of `data.frame` containing the estimated agreements (ppa, npa, ppv, npv) and the lower and upper bounds of the confidence interval for all the methods in `methods_pa` or `methods_pv`.

Author(s)

Lei Yang, Cui Guo, Kate Li, Chang Xu (cuguo@foundationmedicine.com)

References

1. Gart John J and Nam Jun-mo (1988). Approximate interval estimation of the ratio of binomial parameters: a review and corrections for skewness, *Biometrics*, 323-338.

2. Katz DJSM, Baptista J, Azen SP and Pike MC (1978). Obtaining confidence intervals for the risk ratio in cohort studies, *Biometrics*, 469-474.

3. Koopman PAR (1984). Confidence intervals for the ratio of two binomial proportions, *Biometrics*, 513-517.

4. Noether Gottfried E (1957). Two confidence intervals for the ratio of two probabilities and some measures of effectiveness, *Journal of the American Statistical Association*, **52**: 36-45.

See Also

`binom.confint` for different methods to obtain a confidence interval on the binomial probability like PPA and NPA.

Examples

```
agreement(x = 90, y = 10, m = 100, n = 80, prev = 0.3)
agreement(x = 84, y = 0, m = 84, n = 97, prev = 0.096)
```

Index