Package 'ONEST'

October 12, 2022

Title Observers Needed to Evaluate Subjective Tests

Version 0.1.0

Author Gang Han [aut, cre],
Baihong Guo [aut]

Maintainer Gang Han <hangang.true@gmail.com>

URL https://github.com/hangangtrue/ONEST

BugReports https://github.com/hangangtrue/ONEST/issues

Description This ONEST software implements the method of assessing the pathologist agreement in reading PD-L1 assays (Reisenbichler et al. (2020 <doi:10.1038/s41379-020-0544-x>)), to determine the minimum number of evaluators needed to estimate agreement involving a large number of raters. Input to the program should be binary(1/0) pathology data, where "0" may stand for negative and "1" for positive. Additional examples were given using the data from Rimm et al. (2017 <doi:10.1001/jamaoncol.2017.0013>).

Depends R (>= 3.5.0)

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2021-07-26 19:40:02 UTC

R topics documented:

Index   **9**

---

empirical   *Precomputed dataset*

Description

A precomputed dataset only used by the ONEST_vignettes function to create the vignettes.

Usage

```
data('empirical')
```

Details

- columns : different permutations
- rows : agreement percentage

NCCN_22c3_t   *Example dataset5*

Description

NCCN 22c3 tumor data from Rimm et al. (2017) used to illustrate the ONEST method.

Usage

```
data('NCCN_22c3_t')
```

Details

- columns : different raters
- rows : different cases

References

[1] Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA Oncol. 2017;3(8):1051-8, DOI: 10.1001/jamaoncol.2017.0013, PMID: 28278348.

NCCN_sp142 *Example dataset3*

Description

NCCN SP142 stromal data from Rimm et al. (2017) used to illustrate the ONEST method.

Usage

```
data('NCCN_sp142')
```

Details

- columns : different raters
- rows : different cases

References

[1] Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA Oncol. 2017;3(8):1051-8, DOI: 10.1001/jamaoncol.2017.0013, PMID: 28278348.

NCCN_sp142_t *Example dataset4*

Description

NCCN SP142 tumor data from Rimm et al. (2017) used to illustrate the ONEST method.

Usage

```
data('NCCN_sp142_t')
```

Details

- columns : different raters
- rows : different cases

References

[1] Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA Oncol. 2017;3(8):1051-8, DOI: 10.1001/jamaoncol.2017.0013, PMID: 28278348.

ONEST   *The R ONEST package*

Description

The ONEST software implements a statistical method in Reisenbichler et al. (2020[1]), to determine the minimum number of evaluators needed to estimate agreement involving a large number of raters. This method could be utilized by regulatory agencies, such as the FDA, when evaluating agreement levels of a newly proposed subjective laboratory test. Input to the program should be binary(1/0) pathology data, where "0" may stand for negative and "1" for positive. The example datasets in this software are from Rimm et al. (2017[2]) (the SP142 assay), and Reisenbichler et al. 2020. This program can run in R version 3.5.0 and above.

References

[1] Reisenbichler ES, Han G, Bellizzi A, Bossuyt V, Brock J, Cole K, et al. Prospective multi-institutional evaluation of pathologist assessment of PD-L1 assays for patient selection in triple negative breast cancer. Mod Pathol. 2020, DOI: 10.1038/s41379-020-0544-x; PMID: 32300181.

[2] Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA Oncol. 2017;3(8):1051-8, DOI: 10.1001/jamaon-col.2017.0013, PMID: 28278348.

ONEST_inflation_test   *The ONEST_inflation_test*

Description

This function uses the binary (0/1) pathology data to test if the agreement proportion will converge to 0 with an increasing number of observers.

Usage

```
ONEST_inflation_test(Y)
```

Arguments

Y   pathology data as a matrix (missing values are allowed)

Value

p_value: chi-square p-value, a small p-value indicates significant evidence that the observers' agreement will converge to a non-zero proportion.

Examples

```
data("sp263_bin")
ONEST_inflation_test(sp263_bin)
```

ONEST_main *ONEST main function*

Description

This is the ONEST main function taking binary (0/1) pathology data as input. This function utilizes the plotline function to produce the ONEST graphs and estimates.

Usage

```
ONEST_main(data)
```

Arguments

data   a matrix containing the binary pathology data. Each row is the data from one case, and each column is the data from one rater. Missing values are allowed and can be denoted as NA or left blank. If there are n cases and k raters, the input 'data' is a matrix with dimension n by k.

Value

consist_p: a vector of length k-1, indicating proportion of identical reads among a set of pathologists. For example, the first element of "consist_p" is the estimate of agreement percentage for 2 raters. The k-1 th element is the estimate of agreement percentage for k raters.

consist_low: a vector of length k-1, indicating the lower bound of the agreement percentage with 95 percent confidence level corresponding to "consist_p".

diff_consist: a vector of length k-2, indicating the difference between the consist_p. For example, the first element of "diff_consist" is the estimated difference of agreement percentage after increasing from 2 to 3 raters. The k-2 th element is the difference of agreement percentage after increasing from k-1 to k raters.

diff_high: a vector of length k-2, indicating the upper bound of the change of agreement percentage corresponding to "diff_consist" with 95 percent confidence level.

size_case: number of cases n

size_rater: number of raters k

p: the probability of of being rated positive among the proportion of '1-p_plus-p_minus' cases.

p_plus: proportion of the cases rated positive by all raters.

p_minus: proportion of the cases rated negative by all raters.

empirical: a matrix of dimension k-1 by 3, including the empirical estimate of the agreement percentage, and the empirical 95 percent confidence intervals (CI) of the agreement percentage with equal tail probabilities on the two sides. The empirical estimate and CI were calculated by permuting the raters with 1000 random permutations, and using the mean, 2.5th percentile, and 97.5th percentile.

Examples

```
data("sp142_bin")
ONEST_main(sp142_bin)
```

ONEST_vignettes   *ONEST vignettes function*

Description

This function is only used to create vignettes. Although it can get the same results from ON-EST_main(sp142_bin), it uses some precomputed data to decrease the time to get the results. Therefore, it can only be applied to the sp142_bin dataset. Please use ONEST_main instead.

Usage

```
ONEST_vignettes(data,empirical)
```

Arguments

data   a matrix containing the binary pathology data. Each row is the data from one
                      case, and each column is the data from one rater. Missing values are allowed
                      and can be denoted as NA or left blank. If there are n cases and k raters, the
                      input 'data' is a matrix with dimension n by k.

empirical   a matrix containing data used to plot the empirical confidence intervals for the
                      sp142_bin dataset.

Value

consist_p: a vector of length k-1, indicating proportion of identical reads among a set of pathologists. For example, the first element of "consist_p" is the estimate of agreement percentage for 2 raters. The k-1 th element is the estimate of agreement percentage for k raters.

consist_low: a vector of length k-1, indicating the lower bound of the agreement percentage with 95 percent confidence level corresponding to "consist_p".

diff_consist: a vector of length k-2, indicating the difference between the consist_p. For example, the first element of "diff_consist" is the estimated difference of agreement percentage after increasing from 2 to 3 raters. The k-2 th element is the difference of agreement percentage after increasing from k-1 to k raters.

diff_high: a vector of length k-2, indicating the upper bound of the change of agreement percentage corresponding to "diff_consist" with 95 percent confidence level.

size_case: number of cases n

size_rater: number of raters k

p: the probability of of being rated positive among the proportion of '1-p_plus-p_minus' cases.

p_plus: proportion of the cases rated positive by all raters.

p_minus: proportion of the cases rated negative by all raters.

empirical: a matrix of dimension k-1 by 3, including the empirical estimate of the agreement percentage, and the empirical 95 percent confidence intervals (CI) of the agreement percentage with equal tail probabilities on the two sides. The empirical estimate and CI were calculated by permuting the raters with 1000 random permutations, and using the mean, 2.5th percentile, and 97.5th percentile.

Examples

```
data('sp142_bin')
data('empirical')
ONEST_vignettes(sp142_bin,empirical)
```

---

plotline   *Plotline*

Description

This function uses the binary (0/1) pathology data to plot the percentage of agreement.

Usage

```
plotline(path,indi=1,color="red")
```

Arguments

path	pathology data as a matrix (missing values are allowed)
indi	an indicator; if indi=0: then do not make plot; indi=1: then make plot, original scale. Labels on the x and y-axis are on the original scale
color	the color used to draw pictures

Value

A matrix with percentage of agreement from the original scale

Examples

```
data("sp142_bin")
plotline(sp142_bin,1,"red")
```

---

sp142_bin   *Example dataset1*

Description

A pathology reads dataset of triple negative breast cancer in Reisenbichler et al. (2020). A value of 0 means negative and a value of 1 means positive.

Usage

```
data('sp142_bin')
```

Details

- columns : different raters
- rows : different cases

References

[1] Reisenbichler ES, Han G, Bellizzi A, Bossuyt V, Brock J, Cole K, et al. Prospective multi-institutional evaluation of pathologist assessment of PD-L1 assays for patient selection in triple negative breast cancer. Mod Pathol. 2020, DOI: 10.1038/s41379-020-0544-x; PMID: 32300181.

---

sp263_bin   *Example dataset2*

Description

A pathology reads dataset of triple negative breast cancer in Reisenbichler et al. (2020). A value of 0 means negative and a value of 1 means positive.

Usage

```
data('sp263_bin')
```

Details

- columns : different raters
- rows : different cases

References

[1] Reisenbichler ES, Han G, Bellizzi A, Bossuyt V, Brock J, Cole K, et al. Prospective multi-institutional evaluation of pathologist assessment of PD-L1 assays for patient selection in triple negative breast cancer. Mod Pathol. 2020, DOI: 10.1038/s41379-020-0544-x; PMID: 32300181.

Index