Package 'CompareTests'

October 12, 2022

Type Package

Title Correct for Verification Bias in Diagnostic Accuracy & Agreement

Version 1.2

Date 2016-2-6

Author Hormuzd A. Katki and David W. Edelstein

Maintainer Hormuzd Katki <katkih@mail.nih.gov>

Description A standard test is observed on all specimens. We treat the second test (or sampled test) as being conducted on only a stratified sample of specimens. Verification Bias is this situation when the specimens for doing the second (sampled) test is not under investigator control. We treat the total sample as stratified two-phase sampling and use inverse probability weighting. We estimate diagnostic accuracy (category-specific classification probabilities; for binary tests reduces to specificity and sensitivity, and also predictive values) and agreement statistics (percent agreement, percent agreement by category, Kappa (unweighted), Kappa (quadratic weighted) and symmetry tests (reduces to McNemar's test for binary tests)). See: Katki HA, Li Y, Edelstein DW, Castle PE. Estimating the agreement and diagnostic accuracy of two diagnostic tests when one test is conducted on only a subsample of specimens. Stat Med. 2012 Feb 28; 31(5) <doi:10.1002/sim.4422>.

License GPL-3

LazyLoad yes

URL http://dceg.cancer.gov/about/staff-directory/biographies/A-J/katki-hormuzd

NeedsCompilation no

Repository CRAN

Date/Publication 2017-02-06 16:37:03

R topics documented:

CompareTests-package   *Correct for Verification Bias in Diagnostic Accuracy & Agreement*

Description

A standard test is observed on all specimens. We treat the second test (or sampled test) as being conducted on only a stratified sample of specimens. Verification Bias is this situation when the specimens for doing the second (sampled) test is not under investigator control. We treat the total sample as stratified two-phase sampling and use inverse probability weighting. We estimate diagnostic accuracy (category-specific classification probabilities; for binary tests reduces to specificity and sensitivity) and agreement statistics (percent agreement, percent agreement by category, Kappa (unweighted), Kappa (quadratic weighted) and symmetry test (reduces to McNemar's test for binary tests)).

Details

Package:	CompareTests
Type:	Package
Version:	1.1
Date:	2015-06-19
License:	GPL-3
LazyLoad:	yes

You have a dataframe with columns "stdtest" (no NAs allowed; all specimens with NA stdtest results are dropped), "sampledtest" (a gold standard which is NA for some specimens), sampling strata "strata1" "strata2" (values cannot be missing for any specimens). Correct for Verification Bias in the diagnostic and agreement statistics with CompareTests(stdtest,sampledtest,interaction(strata1,strata2),goldstd="sampledtest"

Author(s)

Hormuzd A. Katki and David W. Edelstein

Maintainer: Hormuzd Katki <katkih@mail.nih.gov>

References

Katki HA, Li Y, Edelstein DW, Castle PE. Estimating the agreement and diagnostic accuracy of two diagnostic tests when one test is conducted on only a subsample of specimens. Stat Med. 2012 Feb 28; 31(5): 10.1002/sim.4422.

Examples

```
# Get specimens dataset
data(specimens)
```

```
# Get diagnostic and agreement statistics if sampledtest is the gold standard
CompareTests(specimens$stdtest,specimens$sampledtest,specimens$stratum)

# Get diagnostic and agreement statistics if stdtest is the gold standard
CompareTests(specimens$stdtest,specimens$sampledtest,specimens$stratum,goldstd="stdtest")

# Get agreement statistics if neither test is a gold standard
CompareTests(specimens$stdtest,specimens$sampledtest,specimens$stratum,goldstd=FALSE)
```

CompareTests   *Correct for Verification Bias in Diagnostic Accuracy & Agreement*

Description

A standard test is observed on all specimens. We treat the second test (or sampled test) as being conducted on only a stratified sample of specimens. We treat the total sample as stratified two-phase sampling and use inverse probability weighting. We estimate diagnostic accuracy (category-specific classification probabilities; for binary tests reduces to specificity and sensitivity) and agreement statistics (percent agreement, percent agreement by category, Kappa (unweighted), Kappa (quadratic weighted) and symmetry tests (reduces to McNemar's test for binary tests)).

Usage

```
CompareTests(stdtest, sampledtest, strata = NA, goldstd = "sampledtest")
```

Arguments

stdtest	A vector of standard test results. Any NA test results are dropped from the analysis entirely.
sampledtest	A vector of test results observed only on a sample of specimens. Test results with NA are assumed to no be observed for that specimen
strata	The sampling stratum each specimen belongs to. Set to NA if no sampling or simple random sampling.
goldstd	For outputing diagnostic accuracy statistics, denote if "stdtest" or "sampledtest" is the gold standard. If no gold standard, set to FALSE.

Value

Outputs to screen the estimated contingency table of paired test results, agreement statistics, and diagnostic accuracy statistics.

Returns a list with the following components

Cells	Observed contingency tables of pair test results for each stratum
EstCohort	Weighted contingency table of each pair of test results
Cellvars	Variance of each weighted cell count

Cellcovars	Variance-covariance matrix for each column of weighted cell counts
p0	Percent agreement
Varp0	Variance of percent agreement
AgrCat	Percent agreement by each test category
VarAgrCat	Variance of Percent agreement by each test category
uncondsymm	Symmetry test test statistic
Margincovars	covariance of each pair of margins
Kappa	Kappa (unweighted)
Kappavar	Variance of Kappa
iPV	Each predictive value (for binary tests, NPV and PPV)
VarsiPV	Variance of each predictive value (for binary tests, NPV and PPV)
iCSCP	Each category-specific classification probability (for binary tests, specificity and sensitivity
VarsiCSCP	Variance of each category-specific classification probability (for binary tests, specificity and sensitivity
WeightedKappa	Kappa (quadratic weights)
varWeightedKappa	
	Variance of quadratic-weighted Kappa

Note

Order the categories from least to most severe, for binary (-,+) or (0,1) to make sure that what is output as sensitivity is not the specificity, or that PPV is not reported as NPV.

If you have multiple variables to be crossed to represent the sampling strata, use interaction(), e.g. strata=interaction(strata1,strata2)

Author(s)

Hormuzd A. Katki and David W. Edelstein

References

Katki HA, Li Y, Edelstein DW, Castle PE. Estimating the agreement and diagnostic accuracy of two diagnostic tests when one test is conducted on only a subsample of specimens. Stat Med. 2012 Feb 28; 31(5): 10.1002/sim.4422.

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--or do  help(data=index)  for the standard data sets.

##
# Stat Med Paper 2x2 Chlamydia testing verification bias example
# Note that p for symmetry test is 0.12 not 0.02 as reported in the Stat Med paper
```

```
##

# Convert 2x2 Chlamydia testing table to a dataframe for analysis
# Include NAs for the samples where CTDT test was not conducted (HC2 was conducted on all)
HC2stdtest <- c(rep(1,827),rep(0,4998))
stratum <- HC2stdtest
CTDTsampledtest <- c(               rep(1,800), # 1,1 cell
                                    rep(0,27), # 1,0 cell HC2+ , CTDT-
                                    rep(NA,827-800-27), # HC2+, and no CTDT test done
                                    rep(1,6), # 0,1 cell: HC2-, CTDT+
                                    rep(0,396),# 0,0 cell: HC2- and CTDT-
                                    rep(NA,4998-6-396) # HC2-, no CTDT test done
)
chlamydia <- data.frame(stratum,HC2stdtest,CTDTsampledtest)

# Analysis
temp <- CompareTests(chlamydia$HC2stdtest,
                     chlamydia$CTDTsampledtest,
                     chlamydia$stratum,
                     goldstd="sampledtest"
)


##
# Example analysis of fictitious data example
##
data(specimens)
temp <- CompareTests(specimens$stdtest,
                     specimens$sampledtest,
                     specimens$stratum,
                     goldstd="sampledtest")
## The output is
# The weighted contingency table:
#                           as.factor.stdtest.
# as.factor.sampledtest.    1       2       3       4
#                      1 47.88   7.158   3.322   0.000
#                      2 20.12 104.006  21.861   2.682
#                      3  0.00  10.836  97.494   8.823
#                      4  0.00   0.000   3.322  74.495
#
#
# Agreement Statistics
#
# pct agree and 95% CI: 0.8057 ( 0.7438 0.8555 )
# pct agree by categories and 95% CI
#     est   left   right
# 1 0.6101 0.4501 0.7494
# 2 0.6241 0.5315 0.7083
# 3 0.6693 0.5562 0.7658
# 4 0.8340 0.6340 0.9358
# Kappa and 95% CI: 0.734 ( 0.6509 0.8032 )
# Weighted Kappa (quadratic weights) and 95% CI: 0.8767 ( 0.7107 0.9536 )
# symmetry chi-square: 9.119 p= 0.167
```

```
#
#
#
# Diagnostic Accuracy statistics
#
#        est   left  right
# 1PV 0.7041 0.5422 0.8271
# 2PV 0.8525 0.7362 0.9229
# 3PV 0.7738 0.6547 0.8605
# 4PV 0.8662 0.6928 0.9490
#          est   left  right
# 1CSCP 0.8204 0.6011 0.9327
# 2CSCP 0.6996 0.6169 0.7710
# 3CSCP 0.8322 0.7219 0.9046
# 4CSCP 0.9573 0.5605 0.9975
```

---

fulltable	*fulltable attaches margins and NA/NaN category to the output of table()*

Description

fulltable attaches margins and NA/NaN category to the output of table()

Arguments

same as table()

Value

same as returned from table()

Author(s)

Hormuzd A. Katki

See Also

table

Examples

```
## The function is currently defined as
function (...)
{
  ## Purpose: Add the margins automatically and don't exclude NA/NaN as its own row/column
  ##          and also add row/column titles.  Works for mixed numeric/factor variables.
  ##          For factors, the exclude option won't include the NAs as columns, that's why
  ##          I need to do more work.
```

```
## ----------------------------------------------------------------------
## Arguments: Same as for table()
## ----------------------------------------------------------------------
## Author: Hormuzd Katki, Date:  5 May 2006, 19:45

# This works for purely numeric input, but not for any factors b/c exclude=NULL won't
# include NAs for them.
#   return(addmargins(table(...,exclude=NULL)))

##
# Factors are harder.  I have to reconstruct each factor to include NA as a level
##

# Put everything into a data frame
x <- data.frame(...)

# For each factor (in columns), get the raw levels out, reconstruct to include NAs
# That is, if there are any NAs -- if none, add it as a level anyway
for (i in 1:dim(x)[2]) {
  if ( is.factor(x[,i]) )
    if ( any(is.na(x[,i])) )
      x[,i] <- factor(unclass(x[,i]),labels=c(levels(x[,i]),"NA"),exclude=NULL)
    else
      levels(x[,i]) <- c(levels(x[,i]),"NA")
}

# Make table with margins.  Since NA is a level in each factor, they'll be included
return(addmargins(table(x,exclude=NULL)))

}
```

---

specimens   *Fictitious data on specimens tested by two methods*

Description

stdtest has been done on everyone, and sampledtest has been done on a stratifed subsample of 275 out of 402 specimens (is NA on the other 127 specimens)

Usage

```
data(specimens)
```

Format

A data frame with 402 observations on the following 3 variables.

stratum 6 strata used for sampling

stdtest standard test result available on all specimens

sampledtest new test result available only on stratified subsample

Examples

```
data(specimens)
```

Index