Package 'crossmatch'

June 21, 2024

Type Package

Title The Cross-Match Test

Version 1.4-0

Date 2024-06-18

Description Performs the cross-match test that is an exact,

distribution free test of equality of 2 high dimensional multivariate distributions. The input is a distance matrix and the labels of the two groups to be compared, the output is the number of

cross-matches and a p-value. See Rosenbaum (2005) <doi:10.1111/j.1467-9868.2005.00513.x>.

Imports nbpMatching

Suggests MASS

License GPL-2

LazyLoad yes

NeedsCompilation no

Author Ruth Heller [aut, cph], Dylan Small [aut, cph], Paul Rosenbaum [aut, cph], Marieke Stolte [cre]

Maintainer Marieke Stolte <stolte@statistik.tu-dortmund.de>

Repository CRAN

Date/Publication 2024-06-21 08:00:15 UTC

Contents

	crossmatchdist.									•			 	•			•			•		 •	 						2
	crossmatchtest .		•	• •		•	•	•	•	•	•	• •	 	•	•	•	•	•	•	•	•	 •	 	•	•	•		•	3
Index																													5

crossmatchdist

Description

The exact null distribution of the number of crossmatches for bigN>=4 cases, n>=2 from one type and N-n>=2 from another type.

Usage

crossmatchdist(bigN, n)

Arguments

bigN	The total number of observations
n	The number of cases from one type

Details

bigN is even. Let a1 be the number of cross-matches pairs. Then a2=(n-a1)/2 and a0=bigN/2-(n+a1)/2 are the number of pairs both of one type and the other type respectively.

Value

dist A matrix with rows a0, a1, a2, Pr(A1=a1) and Pr(A1<=a1).

Author(s)

Ruth Heller

References

Rosenbaum, P.R. (2005), An exact distribution-free test comparing two multivariate distributions based on adjacency, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67**, 4, 515-530.

Examples

crossmatchdist(18,9)

Description

A test for comparing two multivariate distributions by using the distance between the observations.

Usage

crossmatchtest(z, D)

Arguments

Z	A binary vector corresponding to observations class labels.
D	A distance matrix of dimensions NxN, where N is the total number of observa-
	tions.

Details

Observations are divided into pairs to minimize the total distance within pairs, using a polynomial time algorithm made available in R by Lu, B., Greevy, R., Xu, X., and Beck, C in the R package "nbpMatching". The cross-match test takes as the test statistic the number of times a subject from one group was paired with a subject from another group, rejecting the hypothesis of equal distribution for small values of the statistic; see Rosenbaum (2005) for details.

Value

A list with the following

a1	The number of cross-matches
Ea1	The expected number of cross-matches under the null
Va1	The variance of number of cross-matches under the null
dev	The observed difference from expectation under null in SE units
pval	The p-value based on exact null distribution (NA for datasets with 340 observations or more)
approxpval	The approximate p-value based on normal approximation

Author(s)

Ruth Heller

References

Rosenbaum, P.R. (2005), An exact distribution-free test comparing two multivariate distributions based on adjacency, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67**, 4, 515-530.

Examples

The example in Section 2 of the article (see References)

```
#The data consists of 2 outcomes measured on 9 treated cases and 9 controls:
dat <- rbind(c(0.47,0.39,0.47,0.78,1,1,0.54,1,0.38,1,0.27,0.63,0.22,0,-1,-0.42,-1,-1),
              c(0.03,0.11,0.16,-0.1,-0.05,0.16,0.12,0.4,0.04,0.71,0.01,0.21,-0.18,
                -0.08, -0.35, 0.26, -0.6, -1.0))
z <- c(rep(0,9),rep(1,9))</pre>
X <- t(dat)
## Rank based Mahalanobis distance between each pair:
X <- as.matrix(X)</pre>
n <- dim(X)[1]
k <- dim(X)[2]</pre>
for (j in 1:k) X[,j] <- rank(X[,j])</pre>
cv <- cov(X)
vuntied <- var(1:n)</pre>
rat <- sqrt(vuntied/diag(cv))</pre>
cv <- diag(rat) %*% cv %*% diag(rat)</pre>
out <- matrix(NA,n,n)</pre>
library(MASS)
icov <- ginv(cv)</pre>
for (i in 1:n) out[i,] <- mahalanobis(X,X[i,],icov,inverted=TRUE)</pre>
dis <- out
## The cross-match test:
crossmatchtest(z,dis)
```

Index

crossmatchdist, 2
crossmatchtest, 3