Package: DOS2 (via r-universe)

October 29, 2024

Type Package

Title Design of Observational Studies, Companion to the Second Edition

Version 0.5.2

Author Paul R. Rosenbaum

Maintainer Paul Rosenbaum <rosenbaum@wharton.upenn.edu>

Description Contains data sets, examples and software from the Second Edition of ``Design of Observational Studies''; see Rosenbaum, P.R. (2010) <doi:10.1007/978-1-4419-1213-8>.

Depends R (>= 3.5.0)

Imports stats, MASS, sensitivity2x2xk, graphics, sensitivitymult, sensitivitymv, senstrat

Suggests optmatch, DiPs

License GPL-2

Encoding UTF-8

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2019-09-16 10:30:05 UTC

Contents

---

addalmostexact   *Use a Penalty to Obtain a Near Exact Match*

Description

A near-exact match (also known as an almost exact match) for a nominal covariate maximizes the number of pairs that are exactly matched for the nominal covariate. This is done by adding a large penalty to the covariate distance between individuals with different values of the nominal covariate. The topic is discussed in Section 10.2 of the second edition of "Design of Observational Studies".

Usage

```
addalmostexact(dmat, z, f, mult = 10)
```

Arguments

dmat	A distance matrix with one row for each treated individual and one column for each control. Often, this is either the Mahalanobis distance based on covariates, 'mahal', or else a robust variant produced by 'smahal'.
z	z is a vector that is 1 for a treated individual and 0 for a control. length(z) must equal sum(dim(dmat)).
f	A vector or factor giving the levels of the nominal covariate. length(f) must equal length(z).
mult	Let mx be the largest distance in dmat. Then mult*mx is added to each distance in dmat for which the treated and control individuals have different values of f. For large enough values of mult, this will ensure that a minimum distance match maximizes the number of individuals who are exactly matched for f.

Details

Very large values of mx may result in numerical instability or slow computations.

Used informally, a small value of mx, say 1 or 1/10, may not maximize the number of exactly matched pairs, but it may usefully increase the number of exactly matched pairs.

Value

Returns the penalized distance matrix.

Note

Section 10.2 of "Design of Observational Studies", second edition, discusses almost-exact matching, also called near-exact matching. The topic is also discussed in section 9.2 of the first edition.

The matching functions in the 'DOS2' package are aids to instruction or self-instruction while reading "Design of Observational Studies". As in the book, these functions break the task of matching into small steps so they are easy to understand in depth. In practice, matching entails a fair amount of book-keeping best done by a package that automates these tasks. Consider R packages 'optmatch', 'rcbalance', 'DiPs', 'designmatch' or 'bigmatch'. Section 14.10 of "Design of Observational Studies", second edition, discusses and compares several R packages for optimal matching.

Author(s)

Paul R. Rosenbaum

References

Zubizarreta, J. R., Reinke, C. E., Kelz, R. R., Silber, J. H. and Rosenbaum, P. R. (2011) <doi:10.1198/tas.2011.11072> "Matching for several sparse nominal variables in a case control study of readmission following surgery". The American Statistician, 65(4), 229-238. Combines near-exact matching with fine balance for the same covariate.

Examples

```
data(costa)
z<-1*(costa$welder=="Y")
aa<-1*(costa$race=="A")
smoker=1*(costa$smoker=="Y")
age<-costa$age
x<-cbind(age,aa,smoker)
dmat<-mahal(z,x)
# Mahalanobis distances
round(dmat[,1:6],2)
# Mahalanobis distanced penalized for mismatching on smoking.
dmat<-addalmostexact(dmat, z, smoker, mult = 10)
# The first treated subject (row labeled 27) is a nonsmoker, but the
# third control (column 3) is a smoker, so there is a big penalty.
round(dmat[,1:6],2)
```

| addcaliper | *Implement a Caliper Using a Penalty Function in Optimal Matching* |

Description

A caliper forces a close match on a continuous covariate, often the propensity score (Rosenbaum and Rubin 1985). Alternatively or additionally, a caliper may be imposed on a prognostic score estimated from some other independent data set (Hansen 2008). Rosenbaum and Rubin (1985) suggested minimizing the Mahalanobis distance within calipers defined by the propensity score. If the caliper is implemented using a penalty function, the caliper cannot create infeasibility. Implementation of a caliper using a penalty function is discussed in Chapter 9 of "Design of Observational Studies", second edition.

Usage

```
addcaliper(dmat, z, p, caliper = 0.1, penalty = 1000)
```

Arguments

dmat	A distance matrix with one row for each treated individual and one column for each control. Often, this is either the Mahalanobis distance based on covariates, 'mahal', or else a robust variant produced by 'smahal'.
z	z is a vector that is 1 for a treated individual and 0 for a control. The number of treated subjects, sum(z), must equal the number of rows of dmat, and the number of potential controls, sum(1-z), must equal the number of columns of dmat.
p	A vector covariate to which the caliper will be applied. Often, p is the propensity score. length(p) must equal length(z).
caliper	A positive number. A penalty will be added to row i and column j of the distance matrix dmat if treated unit i and control unit j have values of p that differ in absolute value by more than caliper*sd(p), the default being a tenth of the standard deviation of p. This is different from the code in the first edition of "Design of Observational Studies", where caliper defaulted to 0.2, rather than 0.1.
penalty	A positive number. penalty determines the magnitude of the penalty that is added to the distance matrix when the caliper is violated. Let pt and pc be the values of p for a treated individual and a control. If |pt-pc| <= caliper*sd(p), then nothing is added to dmat for these individuals. Otherwise, define gap = |pt-pc|-caliper*sd(p); then, the distance between these individuals in dmat is increased by gap*penalty. So there is no penalty inside the caliper, but dmat is increased rapidly as violation of the caliper increases in magnitude. A reasonable value of penalty will depend upon the units in which p is measured, and the default values are reasonable starting points for the propensity score expressed as a probability.

Details

By calling 'addcaliper' several times, calipers may be imposed on several variables, say a propensity score and age. If you use several penalties, say from addcaliper, addalmostexact, and fine, then you need to pay some attention to their relative magnitudes to ensure that they express your view of the relative importance of the conditions they impose.

Value

Returns the penalized distance matrix.

Note

The use of a penalty function to implement a caliper is discussed in chapter 9 of the second edition of "Design of Observational Studies". It is also discussed in chapter 8 of the first edition.

The matching functions in the 'DOS2' package are aids to instruction or self-instruction while reading "Design of Observational Studies". As in the book, these functions break the task of matching into small steps so they are easy to understand in depth. In practice, matching entails a fair amount of book-keeping best done by a package that automates these tasks. Consider R packages 'optmatch', 'rcbalance', 'DiPs', 'designmatch' or 'bigmatch'. Section 14.10 of "Design of Observational Studies", second edition, discusses and compares several R packages for optimal matching.

Author(s)

Paul R. Rosenbaum

References

Hansen, B. B. and Klopfer, S. O. (2006) <doi:10.1198/106186006X137047> "Optimal full matching and related designs via network flows". Journal of Computational and Graphical Statistics, 15(3), 609-627. ('optmatch' package)

Hansen, B. B. (2007) <https://www.r-project.org/conferences/useR-2007/program/presentations/hansen.pdf> "Flexible, optimal matching for observational studies". R News, 7, 18-24. ('optmatch' package)

Hansen, B. B. (2008) <doi:10.1093/biomet/asn004> "The prognostic analogue of the propensity score". Biometrika, 95(2), 481-488.

Rosenbaum, P. R. and Rubin, D. B. (1985) <doi:10.1080/00031305.1985.10479383> "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score". The American Statistician, 39, 33-38.

Examples

```
data(costa)
z<-1*(costa$welder=="Y")
aa<-1*(costa$race=="A")
smoker=1*(costa$smoker=="Y")
age<-costa$age
x<-cbind(age,aa,smoker)
dmat<-mahal(z,x)
# Mahalanobis distances
```

```
round(dmat[,1:6],2)
# Compare with Table 9.5 in Design of Observational
# Studies, 2nd edition.
# Impose propensity score calipers
prop<-glm(z~age+aa+smoker,family=binomial)$fitted.values # propensity score
# Mahalanobis distanced penalized for violations of a propensity score caliper.
# This version is used for numerical work.
dmat<-addcaliper(dmat,z,prop,caliper=.5)
round(dmat[,1:6],2)
# Compare with Table 9.5 in "Design of Observational
# Studies", 2nd edition.
## Not run:
# Find the minimum distance match within propensity score calipers.
optmatch::pairmatch(dmat,data=costa)

## End(Not run)
# Conceptual versions with infinite
# distances for violations of propensity caliper.
dmat[dmat>20]<-Inf
round(dmat[,1:6],2)
# Compare with Table 9.5 in "Design of Observational
# Studies", 2nd edition.
```

---

  angristlavy   *Class Size and Academic Performance – Maimonidies Rule*

Description

This data set is from Angrist and Lavy (1999). There are 86 pairs of two Israeli schools, one with slightly more than 40 students in the fifth grade, the other with 40 or fewer in the fifth grade, together with test scores in reading and math. This example is discussed in Chapters 1, 5 and 18 of "Design of Observational Studies".

Usage

```
data("angristlavy")
```

Format

A data frame with 172 observations on the following 9 variables.

scode School ID

numclass Number of classes in the fifth grade, 1 or 2.

cohsize Total number of students in the fifth grade, near 40 for these schools.

avgmath Average grade in math in the fifth grade.

avgverb Average verbal grade in the fifth grade.

tipuach Percent of disadvantaged students. Used to form matched pairs.

clasz Average class size in the fifth grade, equal to cohsize/numclass

z 1 if cohsize<=40, 0 if cohsize>40.

pair pair ID, 1, 2, ..., 86

Details

This example is discussed in Chapters 1, 5, and 18 of the second edition of "Design of Observational Studies".

As discussed by Angrist and Lavy (1999), Maimonidies rule requires that a class of more than 40 be divided to form two or more classes of at most 40, so there is a large discontinuity in class size at 40: at 40 students in the 5th grade, there is one class of 40, but at 41 students, there are two classes of average size 20.5. So the enrolement of one student should cut the class size roughly in half. Adherence to Maimonidies rule is good but not perfect. Pairs of schools were matched for the percent of disadvantaged students (tipuach).

References

Angrist, J. D. and Lavy, V. (1999) <doi:10.1162/003355399556061> "Using Maimonides' rule to estimate the effect of class size on scholastic achievement". The Quarterly Journal of Economics, 114, 533-575.

Angrist, J. D. and Krueger, A. B. (1999) <doi:10.1016/S1573-4463(99)03004-7> "Empirical strategies in labor economics". In Handbook of Labor Economics (Vol. 3, pp. 1277-1366). Elsevier.

Examples

```
# Figure 1.1 in Chapter 1 of "Design of Observational Studies", 2nd edition
data(angristlavy)
attach(angristlavy)
grp<-factor(z,levels=c(1,0),labels=c("31-40","41-50"),ordered=TRUE)
oldpar<-par(mfrow=c(2,2))
boxplot(tipuach~grp,main="Disadvantaged",ylab="Percent")
boxplot(clasz~grp,main="Class Size",ylab="Students")
boxplot(avgmath~grp,main="Math",ylab="Average Score")
boxplot(avgverb~grp,main="Verbal",ylab="Average Score")
detach(angristlavy)
par(oldpar)
```

---

antineoplastic *Biomonitoring of Workers Exposed to Antineoplastic Drugs*

Description

Two groups of exposed workers compared to unexposed controls, where exposed workers prepared antineoplasitic drugs, protected either only by gloves or by gloves and a laminar hood with vertical air flow. The outcome is the comet assay applied to blood lymphocytes. The comet assay is a measure of damage to DNA. Data from Kopjar and Garaj-Vrhovac (2001). Illustrates the concept of evidence factors in Chapter 20 of Design of Observational Studies, second edition.

Usage

```
data("antineoplastic")
```

Format

A data frame with 59 observations on the following 9 variables.

id ID number. See Tables I and III of Kopjar and Garaj-Vrhovac (2001).

age Age in years

str Three age strata, cut at the thirds of age.

grp Gloves = exposed nurses/doctors protected only by gloves, Hood = exposed nurses/doctors protected by both gloves and a laminar hood with vertical air flow, Control = students and office workers not exposed to antineoplastic drugs.

tailmoment Tail moment of the comet assay.

taillength Tail length of the comet assay mu-m.

z1 1 if exposed, 0 if control

z2 1 if Gloves, 0 if Hood, NA if control

f2 TRUE if in factor 2, FALSE otherwise

Details

The data set is intended to illustrate evidence factors, comparing exposed workers to controls, and workers protected only by gloves to workers with the additional protecting of a laminar hood with vertical air flow..

Source

From Tables I and III of Kopjar and Garaj-Vrhovac (2001).

References

Kopjar, N. and Garaj-Vrhovac, V. (2001) <doi:10.1093/mutage/16.1.71> "Application of the alkaline comet assay in human biomonitoring for genotoxicity: a study on Croation medical personnel handling antineoplastic drugs". Mutagenesis 16, 71-78.

Examples

```
data(antineoplastic)
attach(antineoplastic)
table(str)
table(grp)
oldpar<-par(mfrow=c(1,2))
boxplot(tailmoment~grp,ylab="Tail Moment",
        main="3 Groups",ylim=c(8,20))
boxplot(tailmoment~z1,names=c("Control","Exposed"),
        main="Factor 1",ylab="Tail Moment",ylim=c(8,20))
boxplot(tailmoment[f2]~z2[f2],names=c("Hood","Gloves"),
        ylab="Tail Moment",main="Factor 2",ylim=c(8,20))
```

```
y<-senstrat::hodgeslehmann(tailmoment,z1,str)
# First factor
senstrat::senstrat(y,z1,str,gamma=20)
# Second factor
senstrat::senstrat(y[f2],z2[f2],str[f2],gamma=2.75)
detach(antineoplastic)
par(oldpar)
```

---

cohere	*Sensitivity Analysis for a Coherent Signed Rank Statistic With Multiplicity Correction*

Description

For multiple outcomes in an observation study, computes a weighted combination of signed rank statistics, one for each outcome, and performs either a one-sided randomization test or an analysis of sensitivity to departures from random assignment; see Rosenbaum (1997). Each matched set contains one treated individual and one control. The signed rank statistics can be either Wilcoxon's signed rank statistic or the new U-statistic in Rosenbaum (2011). The Scheffe method is described in Rosenbaum (2016). For one outcome, use the functions 'senWilcox' or 'senU', instead of 'cohere'. The method is discussed in Sections 5.2.3, 18.2 and 18.3 of "Design of Observational Studies", second edition.

Usage

```
cohere(y,z,mpair,w=NULL,gamma=1,m=NULL,m1=NULL,m2=NULL,
                apriori=FALSE,Scheffe=FALSE,exact=NULL)
```

Arguments

y	A matrix of responses with no missing data. Different columns of y are different variables. If present, the column names of y are used to label output.
z	Treatment indicators, z=1 for treated, z=0 for control with length(z)==dim(y)[1].
mpair	Matched set indicators, 1, 2, ..., sum(z) with length(mset)==dim(y)[1]. The vector mset may contain integers or may be a factor.
w	Vector of weights to be applied to the signed rank statistics for the several outcomes with length(w)==dim(y)[2]. At least one weight must be nonzero. If w is NULL, then w=c(1,1,...,1) is used, as in Rosenbaum (1997).
gamma	gamma is the sensitivity parameter Γ, where $\Gamma \geq 1$. Setting $\Gamma = 1$ is equivalent to assuming randomized treatment assignment within the matched pairs, and it performs a randomization test.
m	See m2.
m1	See m2.

m2	The default sets m, m1 and m2 to NULL, and in this case Wilcoxon's signed rank statistic is used. Otherwise, the new U-statistic in Rosenbaum (2011) is used, and the quantities (m,m1,m2) define the U-statistic, as they do for a single outcome in the senU() function. If (m,m1,m2) are three integers such that 1 <= m1 <= m2 <= m, then the triple (m,m1,m2) defines a U statistic. If (m,m1,m2) = (1,1,1), then the U statistic is the sign test statistic. If (m,m1,m2) = (2,2,2), then it is the U statistic that closely approximates Wilcoxons signed rank test. If m=m1=m2, then the U statistic is the test of Stephenson (1981). The general U statistic is discussed in Rosenbaum (2011).
apriori	If Scheffe=FALSE and apriori=TRUE, then the weights w are assumed to have been chosen a priori, and a one-sided, uncorrected P-value is reported for gamma=1 or an upper bound on the one-sided, uncorrected P-value is reported for gamma>1. In either case, this is a Normal approximation based on the central limit theorem and equals 1-pnorm(deviate).
Scheffe	If Scheffe=TRUE, then the weights w are assumed to have been chosen after looking at the data. In this case, the P-value or P-value bound is corrected using Scheffe projections. The approximate corrected P-value or P-value bound is 1-pchisq(max(0,deviate)^2,dim(y)[2]). See Rosenbaum (2016). The Scheffe correction permits you to look at every possible w, controlling the family-wise error rate. In particular, a Scheffe correction entitles you to look in both tails, which you do by considering both w and -w.
	See the planScheffe() function and Rosenbaum (2019) for a combination of an apriori and Scheffe comparisons, as discussed in section 18.3 of "Design of Observational Studies", second edition.
	If Scheffe=FALSE and apriori=FALSE, then the deviate is returned, but no P-value is given. This is useful with planScheffe() because the one planned comparison and the infinitely many discovered comparisons have different critical values.
exact	exact plays no role if m=NULL, m1=NULL, m2=NULL. Otherwise, exact plays the same role in cohere() that it plays in senU(). In the new U-statistic in Rosenbaum (2011), there are both exact ranks and approximate ranks, and exact determines which will be used. Approximate ranks are more appropriate if the sample size is large. The exact ranks use combinatorial coefficients that become very large when the sample size is large. If exact is NULL, then exact is set to TRUE if length(z) <= 50, and is set to FALSE if length(z) > 50. The ranks used by the U statistic involve combinatorial coefficiencts that grow rapidly with increasing sample size. If exact=TRUE, these ranks are computed exactly using expression (8) in Rosenbaum (2011). If exact=FALSE, the ranks are computed by an asymptotic approximation that does not involve large combinatorial coefficients, specifically expression (9) in Rosenbaum (2012).

Details

With w=c(1,1,...,1) and apriori=TRUE, this is the coherent Wilcoxon signed rank statistic in Rosenbaum (1997) and Rosenbaum (2002, Section 9.2). Essentially, the signed rank statistics for several outcomes are added together. It is a one-sided test, in which the several outcomes need to pointed in the same direction, so that large values of each column of y signify the expected direction of the

treatment effect. Setting a non-null value for (m,m1,m2) is the analogous statistic but with the ranks in Rosenbaum (2011) instead of Wilcoxon's ranks.

With Scheffe=TRUE, the procedure permits you to test every w that is not (0,0,...,0), yet control the family-wise error rate. This is essentially the method in Rosenbaum (2016), except the test-statistics are signed rank statistics rather than M-statistics.

Used in conjunction with the planScheffe() function, you can test one planned w and all possible w's controlling the family-wise error rate. The size or alpha for the test is shared equitably between the one planned w and all the other w's, so the critical value for the one planned comparison is not greatly increased. The one-sided critical value for the one planned w is close to the critical value for a two-sided test, although you do test in the opposite tail when you try -w in place of w, albeit with a different critical value. See Rosenbaum (2019).

Value

deviate	The upper bound on the standardized deviate that is used to approximate P-values using the Normal or chi-square distribution; see apriori and Scheffe in the arguments.
aprioriPval	If Scheffe=FALSE and apriori=TRUE, the deviate is compared to the upper tail of the Normal distribution to produce either a P-value for gamma=1 or an upper bound on the P-value for gamma>1.
ScheffePval	If Scheffe=TRUE, the deviate is compared to the upper tail of the chi-square distribution to produce either a P-value for gamma=1 or an upper bound on the P-value for gamma>1.
weights	The weights possibly relabeled with colnames of y.

Note

This method is discussed in chapters 5 and 18 of the second edition of "Design of Observational Studies".

For confidence intervals for individual outcomes, use function senWilcox() or senU().

Author(s)

Paul R. Rosenbaum.

References

Rosenbaum, P. R. (1997)<doi:10.2307/2533957> "Signed rank statistics for coherent predictions". Biometrics, 50, 368-374.

Rosenbaum, P. R. (2002) <doi:10.1007/978-1-4757-3692-2_3> "Observational Studies" (2nd Edition). New York: Springer. See section 9.2.

Rosenbaum, P. R. (2011) <doi:10.1111/j.1541-0420.2010.01535.x> "A new U-statistic with superior design sensitivity in matched observational studies". Biometrics, 67(3), 1017-1027.

Rosenbaum, P. R. (2016) <doi:10.1214/16-AOAS942> "Using Scheffe projections for multiple outcomes in an observational study of smoking and periondontal disease". Annals of Applied Statistics, 10, 1447-1471.

Rosenbaum, P. R. (2019) <doi:10.1093/biostatistics/kxy055> "Combining planned and discovered comparisons in observational studies". Biostatistics, to appear.

Scheffe, H. (1953) <doi:10.1093/biomet/40.1-2.87> "A method for judging all contrasts in the analysis of variance". Biometrika, 40, 87-104.

Examples

```
# Reproduces parts of Table 5.3 from Design of Observational Studies
data(angristlavy)
attach(angristlavy)
cohere(cbind(avgmath,avgverb),1-z,pair,w=c(1,1),apriori=TRUE)
cohere(cbind(avgmath,avgverb),1-z,pair,w=c(1,0),apriori=TRUE)
cohere(cbind(avgmath,avgverb),1-z,pair,w=c(0,1),apriori=TRUE)
cohere(cbind(avgmath,avgverb),1-z,pair,w=c(1,1),gamma=1.65,apriori=TRUE)

# Uses the technique in Rosenbaum (2019)
# Rejection occurs at gamma=1.5 as 2.054>1.895
cohere(cbind(avgmath,avgverb),1-z,pair,w=c(1,1),gamma=1.5)
planScheffe(2)


detach(angristlavy)

data(teeth)
attach(teeth)
# Coherent Wilcoxon signed rank test
cohere(cbind(either4up,either4low),smoker,mset,apriori=TRUE)

# Sensitivity analysis, gamma=2
cohere(cbind(either4up,either4low),smoker,mset,
       gamma=2,apriori=TRUE)

# Upper teeth only
cohere(cbind(either4up,either4low),smoker,mset,
       w=c(1,0),gamma=2,apriori=TRUE)
# This is the same as the univariate test
y<-either4up[smoker==1]-either4up[smoker==0]
senWilcox(y,gamma=2)

# Try various weights, correcting by Scheffe's method
cohere(cbind(either4up,either4low),smoker,mset,
       w=c(1,2),gamma=2,Scheffe=TRUE)

# Replace Wilcoxon's ranks with the new U-statistic
cohere(cbind(either4up,either4low),smoker,mset,
       w=c(1,2),gamma=2,Scheffe=TRUE,m=8,m1=6,m2=8)
detach(teeth)
```

---

costa   *Welding and DNA-Protein Crosslinks*

Description

This data set is from Costa et al. (1993) and it describes 21 welders and 26 potential controls. All are men. The outcome is a measure of genetic damage; specifically, dpc is a measure of DNA-protein cross-links. There are 3 covariates, age, race and smoking. This tiny example is used to illustrate the concepts of multivariate matching in Chapter 9 of "Design of Observational Studies", second edition. The example is useful because its tiny size permits close inspection of the details of multivariate matching, but its small sample size and limited number of covariates make it highly atypical of matching in observational studies.

Usage

```
data("costa")
```

Format

A data frame with 47 observations on the following 6 variables.

subject Within group ID number.

age Age in years.

race AA=African-American, C=Caucasian

smoker Y=yes, N=no

welder Y=yes/treated, N=no/control

dpc DNA-Protein Cross-links (percent)

Source

The data are from Costa et al. (1993). The data are used as a tiny example in Chapter 9 of "Design of Observational Studies", second edition.

References

Costa, M., Zhitkovich, A. and Toniolo, P. (1993) <https://cancerres.aacrjournals.org/content/53/3/460> "DNA-protein cross-links in welders: molecular implications". Cancer research, 53(3), 460-463.

Examples

```
data(costa)
boxplot(costa$dpc~costa$welder,
  xlab="Control (N) or Welder (Y)",
  ylab="DNA-Protein Cross-links Percent")
```

---

crosscut   *Crosscut Test and its Sensitivity Analysis*

Description

Computes the cross-cut test and its sensitivity analysis. The cross-cut test is a nonparametric test of dose-response correlation with good design sensitivity when used for causal inference in observational studies.

Usage

```
crosscut(x, y, ct = 0.25, gamma = 1, LS=FALSE)
```

Arguments

x	Doses of treatment.
y	Response.
ct	The quantile that defines the cross-cut. By default, the cross-cut is at the outer .25 of the data, the lower 25 percent and the upper 75 percent.
gamma	Sensitivity parameter, gamma>=1.
LS	If LS=TRUE, a large sample test is performed. If LS=FALSE, an exact test is performed. For LS=FALSE, the mh function in the 'sensitivity2x2xk' package is used. For LS=TRUE, the mhLS function in the 'sensitivity2x2xk' package is used.

Details

Performs a one-sided test of no association against positive association, together with a sensitivity analysis. The method is described in Rosenbaum (2016), used in Rosenbaum (2017). An adaptive cross-cut statistic is discussed in Rosenbaum and Small (2017); it cuts at several quantiles and picks the best. See Section 19.4 of "Design of Observational Studies"", second edition.

Value

quantiles	Quantiles that define the crosscut
table	A 2x2 table
output	Output from mh or mhLS when applied to table. The functions mh and mhLS are from the sensitivity2x2xk package. The output includes a one-sided P-value.

Note

The 'crosscut' function makes use of 'mh' and 'mhLS' from the 'sensitivity2x2xk' package.

Author(s)

Paul R. Rosenbaum

References

Rosenbaum, P. R. (2016) <doi:10.1111/biom.12373> "The crosscut statistic and its sensitivity to bias in observational studies with ordered doses of treatment". Biometrics, 72(1), 175-183.

Rosenbaum, P. R. (2017) <doi:10.1214/17-STS621> "The general structure of evidence factors in observational studies". Statist Sci 32, 514-530.

Rosenbaum, P. R. and Small, D. S. (2017) <doi:10.1111/biom.12591> "An adaptive Mantel–Haenszel test for sensitivity analysis in observational studies". Biometrics, 73(2), 422-430.

Examples

```
data(periodontal)
attach(periodontal)
crosscut(cigsperday[z==1],pcteither[z==1]-pcteither[z==0],ct=.2)
crosscut(cigsperday[z==1],pcteither[z==1]-pcteither[z==0],ct=.2,gamma=1.25)
crosscut(cigsperday[z==1],pcteither[z==1]-pcteither[z==0],ct=.2,gamma=1.25,LS=TRUE)
crosscut(cigsperday[z==1],pcteither[z==1]-pcteither[z==0],ct=1/3)
detach(periodontal)
```

---

crosscutplot	*Scatterplot Associated with the Cross-Cut Test*

Description

A scatterplot of y against x, with the points used by the cross-cut test in black, and the remaining points in grey. See Figure 19.2 in "Design of Observational Studies", second edition.

Usage

```
crosscutplot(x, y, ct = 0.25, xlab = "", ylab = "", main = "", ylim = NULL)
```

Arguments

x	Variable to be plotted on the horizontal axis.
y	Variable to be plotted on the vertical axis.
ct	The quantile that defines the crosscut. By default, the crosscut is at the outer .25 of the data, the lower 25 percent and the upper 75 percent.
xlab	Label for the x axis.
ylab	Label for the y axis.
main	Title of the plot.
ylim	Limits for the y axis. See example.

Value

A scatter plot.

Author(s)

Paul R. Rosenbaum

References

Rosenbaum, P. R. (2016) <doi:10.1111/biom.12373> "The crosscut statistic and its sensitivity to bias in observational studies with ordered doses of treatment". Biometrics, 72(1), 175-183.

Rosenbaum, P. R. (2017) <doi:10.1214/17-STS621> "The general structure of evidence factors in observational studies". Statistical Science 32, 514-530.

Examples

```
# Figure 1 in Rosenbaum (2017)
data(periodontal)
attach(periodontal)
m<-matrix(1:2,1,2)
layout(m,widths=c(1,2))
boxplot(pcteither[z==1]-pcteither[z==0],ylab="Smoker-Control Difference",
        main="(i)",xlab="Matched Pairs",ylim=c(-100,100))
abline(h=0,lty=2)
crosscutplot(cigsperday[z==1],pcteither[z==1]-pcteither[z==0],
    ct=.2,ylab="Smoker-Control Difference",
    xlab="Cigarettes per Day",main="(ii)",ylim=c(-100,100))
abline(h=0,lty=2)
detach(periodontal)
layout(1)
```

---

dynarski   *A Natural Experiment Concerning Subsidized College Education*

Description

The data are from Susan Dynarski (2003)'s study of the effects of a subsidy for college education provided Social Security to children whose fathers had died. The subsidy was eliminated in 1982. As presented here, the data are only a portion of the data used in Dynarski (2003), specifically the period from 1979-1981 when the subsidy was available. The data set here is Table 14.1 of "Design of Observational Studies" (2nd edition), where it was used with the limited goal of illustrating matching techniques.

Usage

```
data("dynarski")
```

Format

A data frame with 2820 observations on the following 10 variables.

id identification number

zb treatment indicator, zb=1 if 1979-1981 with father deceased, or zb=0 if 1979-1981 with father alive

faminc family income, in units of tens of thousands of dollars

incmiss indicator for missing family income

black 1 if black, 0 otherwise

hisp 1 if hispanic, 0 otherwise

afqtpct test score percentile, Armed Forces Qualifications Test

edmissm indicator for missing education of mother

edm education of mother, 1 is <high school, 2 is high school, 3 is some college, 4 is a BA degree or higher

female 1 if female, 0 if male

Details

The examples reproduce steps in Chapter 14 of "Design of Observational Studies" (2nd edition). Portions of the examples require you to load Ben Hansen's 'optmatch' package and accept its academic license; these portions of the examples do not run automatically. Hansen's 'optmatch' package uses the Fortran code of Bertsekas and Tseng (1988).

Source

Dynarski (2003).

References

Bertsekas, D. P. and Tseng, P. (1988) <doi:10.1007/BF02288322> "The relax codes for linear minimum cost network flow problems". Annals of Operations Research, 13(1), 125-190.

Dynarski, S. M. (2003) <doi:10.1257/000282803321455287> "Does aid matter? Measuring the effect of student aid on college attendance and completion". American Economic Review, 93(1), 279-288.

Hansen, B. B. (2007) <https://www.r-project.org/conferences/useR-2007/program/presentations/hansen.pdf> "Flexible, optimal matching for observational studies". R News, 7, 18-24.

Hansen, B. B. and Klopfer, S. O. (2006) <doi:10.1198/106186006X137047> "Optimal full matching and related designs via network flows". Journal of Computational and Graphical Statistics, 15(3), 609-627.

Rosenbaum, P. R. (1989). "Optimal matching for observational studies" <doi:10.1080/01621459.1989.10478868> Journal of the American Statistical Association, 84(408), 1024-1032.

Rosenbaum, P. R. (1991) <doi:10.1111/j.2517-6161.1991.tb01848.x> A characterization of optimal designs for observational studies. Journal of the Royal Statistical Society: Series B (Methodological), 53(3), 597-610.

Examples

```
#
data(dynarski)
# Table 14.1 of "Design of Observational Studies" (2nd edition)
head(dynarski)
# Table 14.2 of "Design of Observational Studies" (2nd edition)
zb<-dynarski$zb
zbf<-factor(zb,levels=c(1,0),labels=c("Father deceased",
    "Father not deceased"))
table(zbf)
Xb<-dynarski[,3:10]

# Estimate the propensity score, Rosenbaum (2010, Section 14.3)
p<-glm(zb~Xb$faminc+Xb$incmiss+Xb$black+Xb$hisp
    +Xb$afqtpct+Xb$edmissm+Xb$edm+Xb$female,
    family=binomial)$fitted.values
# Figure 14.1 in "Design of Observational Studies" (2nd edition)
boxplot(p~zbf,ylab="Propensity score",main="1979-1981 Cohort")

# Read about missing covariate values in section 14.4
# of "Design of Observational Studies" (2nd edition)

# Robust Mahalanobis distance matrix, treated x control
dmat<-smahal(zb,Xb)
dim(dmat)
# Table 14.3 in "Design of Observational Studies" (2nd edition)
round(dmat[1:5,1:5],2)

# Add a caliper on the propensity score using a penalty function
dmat<-addcaliper(dmat,zb,p,caliper=.2)
dim(dmat)
# Table 14.4 in "Design of Observational Studies" (2nd edition)
round(dmat[1:5,1:5],2)
## Not run:
# YOU MUST LOAD the 'optmatch' package and accept its license to continue.
# Note that the 'optmatch' package has changed since 2010.  It now suggests
# that you indicate the data explicitly as data=dynarski.

# Creating a 1-to-10 match, as in section 14.6 of
# "Design of Observational Studies" (2nd edition)
# This may take a few minutes.
m<-fullmatch(dmat,data=dynarski,min.controls = 10,max.controls = 10,
    omit.fraction = 1379/2689)
length(m)
sum(matched(m))
1441/11 # There are 131 matched sets, 1 treated, 10 controls

# Alternative, simpler code to do the same thing
m2<-pairmatch(dmat,controls=10,data=dynarski)
# Results are the same:
sum(m[matched(m)]!=m2[matched(m2)])
```

```
# Housekeeping
im<-as.integer(m)
dynarski<-cbind(dynarski,im)
dm<-dynarski[matched(m),]
dm<-dm[order(dm$im,1-dm$zb),]

# Table 14.5 in "Design of Observational Studies" (2nd edition)
which(dm$id==10)
dm[188:198,]
which(dm$id==396)
dm[23:33,]
which(dm$id==3051)
dm[1068:1078,]
# In principle, there can be a tie, in which several different
# matched samples all minimize the total distance.  On my
# computer, this calculation reproduces Table 14.5, but were
# there a tie, 'optmatch' should return one of the tied optimal
# matches, but not any particular one.

## End(Not run)
```

---

fine   *Expand a Distance Matrix for Matching with Fine Balance.*

Description

In optimal pair matching with fine balance, expand a distance matrix to become a square matrix to enforce fine balance. The method is discussed in Chapter 11 of "Design of Observational Studies", second edition, and it is conceptually the simplest way to implement fine balance; therefore, it remains very useful for teaching and for self-study. See details.

Usage

```
fine(dmat, z, f, mult = 100)
```

Arguments

dmat
: A distance matrix with one row for each treated individual and one column for each control. Often, this is either the Mahalanobis distance based on covariates, 'mahal', or else a robust variant produced by 'smahal'. The distance matrix dmat may have been penalized by 'addalmostexact' or 'addcaliper'. An error will result unless dmat has more columns than rows.

z
: z is a vector that is 1 for a treated individual and 0 for a control. The number of treated subjects, sum(z), must equal the number of rows of dmat, and the number of potential controls, sum(1-z), must equal the number of columns of dmat.

f
: A factor or vector to be finely balanced. Must have length(f)=length(z).

mult	A positive number, mult>0. Determines the penalty used to enforce fine balance as max(dmat)*mult. The distance matrix dmat may have been penalized by 'addalmostexact' or 'addcaliper', and in this case it makes sense to set mult=1 or mult=2, rather than the default, mult=100. If dmat is already penalized, taking mult>1 creates a larger penalty for deviations from fine balance than the exisiting penalties.

Details

The method is discussed in Chapter 11 of "Design of Observational Studies", second edition, and it is conceptually the simplest way to implement fine balance. However, the expanded distance matrix can become quite large, so this simplest method is not the most efficient method in its use of computer storage. A more compact implementation uses minimum cost flow in a network (Rosenbaum 1989, Section 3.2). Additionally, there are several extensions of fine balance, including near-fine balance (Yang et al. 2012, in Yu's 'DiPs' package), fine balance for several covariates via integer programming (Zubizarreta 2012, 'designmatch' R-package), and refined balance (Pimentel et al. 2015, 'rcbalance' R-package). Ruoqi Yu's 'bigmatch' R-package implements fine balance and near-fine balance in very large matching problems.

Value

An expanded, square distance matrix with "extra" treated units for use in optimal pair matching. Any control paired with an "extra" treated unit is discarded, as are the "extra" treated units.

Note

Fine balance is discussed in chapter 11 of "Design of Observational Studie", second edition.

The matching functions in the 'DOS2' package are aids to instruction or self-instruction while reading "Design of Observational Studies". As in the book, these functions break the task of matching into small steps so they are easy to understand in depth. In practice, matching entails a fair amount of book-keeping best done by a package that automates these tasks. For fine balance and similar methods, use R packages 'rcbalance', 'DiPs', 'designmatch' or 'bigmatch'. Section 14.10 of "Design of Observational Studies", second edition, discusses and compares several R packages for optimal matching.

Author(s)

Paul R. Rosenbaum

References

Hansen, B. B. and Klopfer, S. O. (2006) <doi:10.1198/106186006X137047> "Optimal full matching and related designs via network flows". Journal of Computational and Graphical Statistics, 15(3), 609-627. The method implemented in Hansen's 'optmatch' package.

Hansen, B. B. (2007) <https://www.r-project.org/conferences/useR-2007/program/presentations/hansen.pdf> "Flexible, optimal matching for observational studies". R News, 7, 18-24. Discusses Hansen's 'optmatch' package.

Pimentel, S. D., Kelz, R. R., Silber, J. H. and Rosenbaum, P. R. (2015) <doi:10.1080/01621459.2014.997879> "Large, sparse optimal matching with refined covariate balance in an observational study of the

health outcomes produced by new surgeons". Journal of the American Statistical Association, 110, 515-527. Introduces an extension of fine balance called refined balance that is implemented in Pimentel's package 'rcbalance'.

Pimentel, S. D. (2016) "Large, Sparse Optimal Matching with R Package rcbalance" <https://obsstudies.org/large-sparse-optimal-matching-with-r-package-rcbalance/> Observational Studies, 2, 4-23. Discusses and illustrates the use of Pimentel's 'rcbalance' package.

Rosenbaum, P. R. (1989). "Optimal matching for observational studies" <doi:10.1080/01621459.1989.10478868> Journal of the American Statistical Association, 84(408), 1024-1032. Discusses and illustrates fine balance using minimum cost flow in a network in section 3.2.

Rosenbaum, P. R., Ross, R. N. and Silber, J. H. (2007) <doi:10.1198/016214506000001059> "Minimum distance matched sampling with fine balance in an observational study of treatment for ovarian cancer". Journal of the American Statistical Association, 102, 75-83. Discusses and illustrates fine balance using optimal assignment.

Yang, D., Small, D. S., Silber, J. H. and Rosenbaum, P. R. (2012) <doi:10.1111/j.1541-0420.2011.01691.x> "Optimal matching with minimal deviation from fine balance in a study of obesity and surgical outcomes". Biometrics, 68, 628-636. Extension of fine balance useful when fine balance is infeasible. Comes as close as possible to fine balance. Implemented as part of the 'rcbalance' and 'DiPs' packages.

Yu, Ruoqi and Rosenbaum, P.R., (2019a) <doi:10.1111/biom.13098> "Directional penalties for optimal matching in observational studies". Biometrics, to appear, Describes the method in Yu's 'DiPs' package.

Yu, Ruoqi, Silber, J.H. and Rosenbaum, P.R. (2019b) <https://www.imstat.org/journals-and-publications/statistical-science/statistical-science-future-papers/> "Matching methods for observational studies derived from large administrative databases". Stat Sci., to appear. Describes the method in Yu's 'bigmatch' package.

Zubizarreta, J. R., Reinke, C. E., Kelz, R. R., Silber, J. H. and Rosenbaum, P. R. (2011) <doi:10.1198/tas.2011.11072> "Matching for several sparse nominal variables in a case-control study of readmission following surgery". The American Statistician, 65(4), 229-238. Combines near-exact matching with fine balance for the same covariate.

Zubizarreta, J. R. (2012) <doi:10.1080/01621459.2012.703874> "Using mixed integer programming for matching in an observational study of kidney failure after surgery". Journal of the American Statistical Association, 107, 1360-1371. Extends the concept of fine balance using integer programming. Implemented in R in the 'designmatch' package.

Examples

```
data(costa)
z<-1*(costa$welder=="Y")
aa<-1*(costa$race=="A")
smoker=1*(costa$smoker=="Y")
age<-costa$age
x<-cbind(age,aa,smoker)
dmat<-mahal(z,x)
# Mahalanobis distances
round(dmat[,1:6],2) # Compare with Table 9.5 in "Design of Observational Studies"
# Impose propensity score calipers
prop<-glm(z~age+aa+smoker,family=binomial)$fitted.values # propensity score
```

```
# Mahalanobis distanced penalized for violations of a propensity score caliper.
# This version is used for numerical work.
dmat<-addcaliper(dmat,z,prop,caliper=.5)
round(dmat[,1:6],2) # Compare with Table 9.5 in "Design of Observational Studies"
# Because dmat already contains large penalties, we set mult=1.
dmat<-fine(dmat,z,aa,mult=1)
dmat[,1:6] # Compare with Table 11.1 in "Design of Observational Studies"
dim(dmat) # dmat has been expanded to be square by adding 5 extras, here numbered 48:52
# Any control matched to an extra is discarded.
## Not run:
# Find the minimum distance match within propensity score calipers.
optmatch::pairmatch(dmat)
# Any control matched to an extra is discarded.  For instance, the optimal match paired
# extra row 48 with the real control in column 7 to form matched set 1.22, so that control
# is not part of the matched sample.  The harmless warning message from pairmatch
# reflects the divergence between the costa data.frame and expanded distance matrix.

## End(Not run)
# Conceptual versions with infinite distances for violations of propensity caliper.
dmat[dmat>20]<-Inf
round(dmat[,1:6],2) # Compare with Table 11.1 in "Design of Observational Studies"
```

| frontseat | *Safety Belts in Vehicle Crashes* |

Description

Data from the US Fatality Analysis Reporting System (FARS) in 2010-2011, as discussed in Rosenbaum (2015) and in the "Design of Observational Studies", second edition, Chapter 7. The data concern crashes in which a driver and a right front passenger were present, following Evans (1986). The data compare the injuries of the driver and passenger, and are particularly interesting when their safety belt use is different. The example illustrates the analysis of a counterclaim.

Usage

```
data("frontseat")
```

Format

A data frame with 17014 observations on the following 7 variables.

restraint Saftey belt use by (Driver,Passenger), where n=unbelted, ls=lap-shoulder belt. A factor with levels ls.ls ls.n n.ls n.n. Here, ls.n means the driver was belted and the passenger was not.

injury Injury of (Driver,Passenger).

injurydif Difference in injury scores, driver-minus-passenger, from -4 to 4. A score of -4 means the driver was uninjured, but the passenger died.

ejection Ejection from the vehicle of the (Driver,Passenger).

ejectiondif 1 if the driver was ejected but the passenger was not, -1 if the passenger was ejected but the driver was not, 0 if their fates were the same.

gender Genders of the (Driver,Passenger).

agedif Difference in ages, driver-minus-passenger.

Details

This example is discussed in "Design of Observational Studies", second edition, Chapter 7.

Details are given in Rosenbaum (2015). A crash, perhaps involving several vehicles, is recorded in FARS only if there is at least one fatality, creating issues of ascertainment (Fisher 1934) that do not affect tests of the hypothesis of no effect, but that do affect estimation. Only tests of no effect are considered in this example.

Source

Rosenbaum (2015)

References

Evans, L. (1986) <doi:10.1016/0001-4575(86)90007-2> "The Effectiveness of Safety Belts in Preventing Fatalities". Accident Analysis and Prevention, 18, 229–241.

Fisher, R.A. (1934) <doi:10.1111/j.1469-1809.1934.tb02105.x> "The Effect of Methods of Ascertainment Upon the Estimation of Frequencies". Annals of Eugenics, 6(1), 13-25.

Imai, K., Keele, L., Yamamoto, T. (2010) <doi:10.1214/10-STS321> "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects". Statistical Science, 25, 51–71.

Rosenbaum, P.R. (2015) <doi:10.1080/01621459.2015.1054489> "Some Counterclaims Undermine Themselves in Observational Studies". Journal of the American Statistical Association, 110:512, 1389-1398.

Examples

```
data(frontseat)
attach(frontseat)
use<-(!is.na(injurydif))
# Compare with Table 1 in Rosenbaum (2015), case ls.n
table(restraint[use])
use<-use&(restraint=="ls.n")
2*sensitivitymv::senmv(-injurydif[use],gamma=5,
      trim=1,lambda=.99)$pval
2*sensitivitymv::senmv(-injurydif[use],gamma=5.5,
      trim=1,lambda=.99)$pval
2*sensitivitymv::senmv(-injurydif[use],gamma=6,
      trim=1,lambda=.99,inner=.25)$pval
2*sensitivitymv::senmv(-injurydif[use],gamma=6.5,
      trim=1,lambda=.99,inner=.25)$pval

# Counterclaim analysis, one ejected individual
# Compare with Table 2 in Rosenbaum (2015), case ls.n
table(ejection,ejectiondif)
```

```
use<-use&(!is.na(ejectiondif))&(ejectiondif!=0)
sum(use)
2*sensitivitymv::senmv(-injurydif[use],gamma=9,
      trim=1,lambda=.99)$pval
2*sensitivitymv::senmv(-injurydif[use],gamma=11,
      trim=1,lambda=.99,inner=.25)$pval
detach(frontseat)
```

---

hcyst   *Homocysteine and Smoking*

Description

NHANES 2005-2006 data on smoking and homocysteine levels in adults. Matched triples, one daily smoker matched to two never smokers in 548 matched sets of size 3.

Usage

```
data("hcyst")
```

Format

A data frame with 1644 observations on the following 13 variables.

SEQN NHANES identification number

z Smoking status, 1 = daily smoker, 0 = never smoker

female 1 = female, 0 = male

age Age in years, >=20, capped at 85

black 1=black race, 0=other

education Level of education, 3=HS or equivalent

povertyr Ratio of family income to the poverty level, capped at 5 times poverty

bmi BMI or body-mass-index

cigsperday30 Cigarettes smoked per day, 0 for never smokers

cotinine Blood cotinine level, a biomarker of recent exposure to tobacco. Serum cotinine in ng/mL

homocysteine Level of homocysteine. Total homocysteine in blood plasma in umol/L

p An estimated propensity score

mset Matched set indicator, 1, 1, 1, 2, 2, 2, ..., 548, 548, 548

Details

The matching controlled for female, age, black, education, income, and BMI. The outcome is homocysteine. The treatment is daily smoking versus never smoking.

Bazzano et al. (2003) noted higher homocysteine levels in smokers than in nonsmokers. The NHANES data have been used several times to illustrate statistical methods; see Pimental et al (2016), Rosenbaum (2018), Yu and Rosenbaum (2019).

This is Match 4 in Yu and Rosenbaum (2019). The match used a propensity score estimated from these covariates, and it was finely balanced for education. It used a rank-based Mahalanobis distance, directional penalties, and a constraint on the maximum distance. The larger data set before matching is contained in the DiPs package as nh0506homocysteine.

The two controls for each smoker have been randomly ordered, so that, for illustration, a matched pair analysis using just the first control may be compared with an analysis of a 2-1 match. See Rosenbaum (2013) and Rosenbaum (2017b, p222-223) for discussion of multiple controls and design sensitivity.

Source

From the NHANES web page, for NHANES 2005-2006. US National Health and Nutrition Examination Survey, 2005-2006. From the US National Center for Health Statistics.

References

Bazzano, L. A., He, J., Muntner, P., Vupputuri, S. and Whelton, P. K. (2003) <doi:10.7326/0003-4819-138-11-200306030-00010> "Relationship between cigarette smoking and novel risk factors for cardiovascular disease in the United States". Annals of Internal Medicine, 138, 891-897.

Pimentel, S. D., Small, D. S. and Rosenbaum, P. R. (2016) <doi:10.1080/01621459.2015.1076342> "Constructed second control groups and attenuation of unmeasured biases". Journal of the American Statistical Association, 111, 1157-1167.

Rosenbaum, P. R. (2007) <doi:10.1111/j.1541-0420.2006.00717.x> "Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies". Biometrics, 2007, 63, 456-464.

Rosenbaum, P. R. and Silber, J. H. (2009) <doi:10.1198/jasa.2009.tm08470> "Amplification of sensitivity analysis in observational studies". Journal of the American Statistical Association, 104, 1398-1405.

Rosenbaum, P. R. (2013) <doi:10.1111/j.1541-0420.2012.01821.x> "Impact of multiple matched controls on design sensitivity in observational studies". Biometrics, 2013, 69, 118-127.

Rosenbaum, P. R. (2015) <https://obsstudies.org/two-r-packages-for-sensitivity-analysis-in-observational-studies/> "Two R packages for sensitivity analysis in observational studies". Observational Studies, v. 1.

Rosenbaum, P. R. (2016) <doi:10.1111/biom.12373> "The crosscut statistic and its sensitivity to bias in observational studies with ordered doses of treatment". Biometrics, 72(1), 175-183.

Rosenbaum, P. R. (2017a) <doi.org/10.1214/18-AOAS1153> "Sensitivity analysis for stratified comparisons in an observational study of the effect of smoking on homocysteine levels". Annals of Applied Statistics, 12, 2312–2334

Rosenbaum, P. R. (2017b) <https://www.hup.harvard.edu/catalog.php?isbn=9780674975576> "Observation and Experiment: An Introduction to Causal Inference". Cambridge, MA: Harvard Univeristy Press, pp 222-223.

Yu, Ruoqi. and Rosenbaum, P. R. (2019) <doi.org/10.1111/biom.13098> "Directional penalties for optimal matching in observational studies". Biometrics, to appear. See Yu's 'DiPs' package.

Examples

```
data(hcyst)
attach(hcyst)
tapply(female,z,mean)
tapply(age,z,mean)
tapply(black,z,mean)
tapply(education,z,mean)
table(z,education)
tapply(povertyr,z,mean)
tapply(bmi,z,mean)
tapply(p,z,mean)
ind<-rep(1:3,548)
hcyst<-cbind(hcyst,ind)
hcystpair<-hcyst[ind!=3,]
rm(ind)
detach(hcyst)

# Analysis of paired data, excluding second control
attach(hcystpair)
y<-log(homocysteine)[z==1]-log(homocysteine)[z==0]
x<-cotinine[z==1]-cotinine[z==0]


senWilcox(y,gamma=1)
senWilcox(y,gamma=1.53)
senU(y,m1=4,m2=5,m=5,gamma=1.53)
senU(y,m1=7,m2=8,m=8,gamma=1.53)
senU(y,m1=7,m2=8,m=8,gamma=1.7)
# Interpretation/amplification of gamma=1.53 and gamma=1.7
# See Rosenbaum and Silber (2009)
sensitivitymult::amplify(1.53,c(2,3))
sensitivitymult::amplify(1.7,c(2,3))


crosscutplot(x,y,ylab="Difference in log(homocysteine)",
   xlab="Difference in Cotinine, Smoker-minus-Control",main="Homocysteine and Smoking")
text(600,1.8,"n=41")
text(600,-1,"n=31")
text(-500,-1,"n=43")
text(-500,1.8,"n=21")

crosscut(x,y)
crosscut(x,y,gamma=1.25)
```

```
# Comparison of pairs and matched triples
# Triples increase power and design sensitivity; see Rosenbaum (2013)
# and Rosenbaum (2017b, p222-223)
library(sensitivitymult)
sensitivitymult::senm(log(homocysteine),z,mset,gamma=1.75)$pval
detach(hcystpair)
attach(hcyst)
sensitivitymult::senm(log(homocysteine),z,mset,gamma=1.75)$pval
# Inner trimming improves design sensitivity; see Rosenbaum (2013)
sensitivitymult::senm(log(homocysteine),z,mset,inner=.5,gamma=1.75)$pval

# Interpretation/amplification of gamma = 1.75
# See Rosenbaum and Silber (2009)
sensitivitymult::amplify(1.75,c(2,3,10))

# Confidence interval and point estimate
# sensitivitymult::senmCI(log(homocysteine),z,mset,inner=.5,gamma=1.5)
detach(hcyst)
```

---

lead   *Lead in Children*

Description

Data from Morton et al. (1982) concerning exposed children whose fathers worked in a battery plant where lead was used in the manufacture of batteries. Exposed children were matched to controls for age and neighborhood. For exposed children, also given are the father's level of exposure to lead at work (level) and the father's hygiene upon leaving the battery plant at the end of the day. This example is used in Chapters 7 and 9 of Rosenbaum (2017).

Usage

```
data("lead")
```

Format

A data frame with 33 observations on the following 6 variables.

control Blood lead level for the control, micrograms of lead per decaliter of whole blood.

exposed Blood lead level for the exposed/treated child, micrograms of lead per decaliter of whole blood.

level Father's level of exposure to lead: a factor with levels high low medium

hyg Father's hygiene before going home: a factor with levels good mod poor

both A factor built from level and hyg: a factor with levels high.ok high.poor low medium

dif Exposed-minus-control pair difference in blood lead levels.

Details

The data were assembled from two published tables in Morton et al. (1982). One matched pair with no control is omitted here. Small ambiguities in assembling a complete data set from two tables were resolved by a throw of the dice; however, it is a reasonable example to illustrate statistical methods. One table described the exposed-versus-control matched pairs, and these are as in the paper. The second table described the exposed individuals, their level of exposure and their hygiene, and again these are as in the paper. The two tables were linked using the blood lead level of the exposed children, and a couple of ties in these lead levels made for small ambiguities about which control responses belong with which hygienes and exposure levels for the exposed children. See, for instance, rows 18 and 22, where both exposed children have blood lead level 34.

Source

Data are from Morton et al. (1982). They were used as an example in Rosenbaum (1991, 2002, 2011, 2017).

References

Morton, D. E., Saah, A. J., Silberg, S. L., Owens, W. L., Roberts, M. A., & Saah, M. D. (1982) <doi:10.1093/oxfordjournals.aje.a113336> "Lead absorption in children of employees in a lead-related industry". American Journal of Epidemiology, 115(4), 549-555.

Rosenbaum, P. R. (1991) <doi:10.1214/aos/1176348141> "Some poset statistics". The Annals of Statistics, 19(2), 1091-1097.

Rosenbaum, P. R. (2002) <doi:10.1007/978-1-4757-3692-2_3> "Observational Studies" (2nd edition). New York: Springer. Section 4.3.

Rosenbaum, P. R. (2011) <doi:10.1111/j.1541-0420.2010.01535.x> "A new U statistic with superior design sensitivity in matched observational studies". Biometrics, 67(3), 1017-1027.

Rosenbaum, P. R. (2017) <https://www.hup.harvard.edu/catalog.php?isbn=9780674975576> "Observation and Experiment: An Introduction to Causal Inference". Cambridge, MA: Harvard University Press. Chapters 7 and 9.

Examples

```
data(lead)
# Reproduces parts of Table 2 in Rosenbaum (2011)
senU(lead$dif,gamma=5.8,m=8,m1=5,m2=8)
senU(lead$dif,gamma=5,m=5,m1=4,m2=5)

# m=2, m1=2, m2=2 is the U-statistic that closely
# resembles Wilcoxon's signed rank test.  Note
# that the results are almost the same.
senWilcox(lead$dif,gamma=5) # In Table 2
senU(lead$dif,gamma=5,m=2,m1=2,m2=2)
```

---

mahal   *Mahalanobis Distance Matrix for Optimal Matching*

Description

Computes a Mahalanobis distance matrix between treated individuals and potential controls; see Rubin (1980) and Rosenbaum and Rubin (1985). The method is discussed in Section 9.3 of "Design of Observational Studies", second edition.

Usage

```
mahal(z, X)
```

Arguments

z   z is a vector that is 1 for a treated individual and 0 for a control.

X   A matrix of continuous or binary covariates. The number of rows of X must equal the length of z.

Value

The distance matrix has one row for each treated individual (z=1) and one column for each potential control (z=0). The row and column names of the distance matrix refer to the position in z, 1, 2, ..., length(z).

Note

The method is discussed in Section 9.3 of "Design of Observational Studies", second edition.

The matching functions in the 'DOS2' package are aids to instruction or self-instruction while reading "Design of Observational Studies". As in the book, these functions break the task of matching into small steps so they are easy to understand in depth. In practice, matching entails a fair amount of book-keeping best done by a package that automates these tasks. Consider R packages 'optmatch', 'rcbalance', 'DiPs', 'designmatch' or 'bigmatch'. Section 14.10 of "Design of Observational Studies", second edition, discusses and compares several R packages for optimal matching.

Author(s)

Paul R. Rosenbaum

References

Hansen, B. B. and Klopfer, S. O. (2006) <doi:10.1198/106186006X137047> "Optimal full matching and related designs via network flows". Journal of computational and Graphical Statistics, 15(3), 609-627. ('optmatch' package)

Hansen, B. B. (2007) <https://www.r-project.org/conferences/useR-2007/program/presentations/hansen.pdf> "Flexible, optimal matching for observational studies". R News, 7, 18-24. ('optmatch' package)

Rosenbaum, P. R. and Rubin, D. B. (1985) <doi:10.1080/00031305.1985.10479383> "Construct-ing a control group using multivariate matched sampling methods that incorporate the propensity score". The American Statistician, 39, 33-38.

Rubin, D. B. (1980) <doi:10.2307/2529981> "Bias reduction using Mahalanobis-metric matching". Biometrics, 36, 293-298.

Examples

```
data(costa)
z<-1*(costa$welder=="Y")
aa<-1*(costa$race=="A")
smoker=1*(costa$smoker=="Y")
age<-costa$age
x<-cbind(age,aa,smoker)
dmat<-mahal(z,x)
# Mahalanobis distances
round(dmat[,1:6],2)
# Compare with Table 9.5 in "Design of Observational Studies", 2nd ed.
# Impose propensity score calipers
prop<-glm(z~age+aa+smoker,family=binomial)$fitted.values # propensity score
# Mahalanobis distanced penalized for violations of a propensity score caliper.
# This version is used for numerical work.
dmat<-addcaliper(dmat,z,prop,caliper=.5)
round(dmat[,1:6],2)
# Compare with Table 9.5 in "Design of Observational Studies", 2nd ed.
## Not run:
# Find the minimum distance match within propensity score calipers.
# You must load the 'optmatch' package to try this example
optmatch::pairmatch(dmat,data=costa)

## End(Not run)
# Conceptual versions with infinite distances for violations of propensity caliper.
dmat[dmat>20]<-Inf
round(dmat[,1:6],2)
# Compare with Table 9.5 in "Design of Observational Studies", 2nd ed.
```

NSW   *National Supported Work Randomized Experiment*

Description

National Supported Work (NSW) randomized experiment in 185 matched pairs. Used as an example in Chapter 2 of Design of Observational Studies.

Usage

```
data("NSW")
```

Format

A data frame with 370 observations on the following 11 variables.

id Matched pair id, 1, 1, 2, 2, ..., 185, 185.

z z=1 for treated, z=0 for control

age Age in years

edu Education in years

black 1=black, 0=other

hisp 1=Hispanic, 0=other

married 1=married, 0=other

nodegree 1=no High School degree, 0=other

re74 Earnings in 1974, a covariate

re75 Earnings in 1975, a covariate

re78 Earnings in 1978, an outcome

Details

Compare with Table 2.2 of Design of Observational Studies.

Source

Dehejia and Wahba (1999).

References

Couch, K.A. (1992) <doi:10.1086/298292> "New evidence on the long-term effects of employment training programs". Journal of Labor Economics 10, 380-388.

Dehejia, R.H., Wahba, W. (1999) <doi:10.1080/01621459.1999.10473858> "Causal effects in non-experimental studies: Reevaluating the evaluation of training programs". Journal of the American Statistical Association 94, 1053-1062.

LaLonde, R.J. (1986) <https://www.jstor.org/stable/1806062> "Evaluating the econometric evaluations of training programs with experimental data". American Economic Review 76, 604-620.

Examples

```
data(NSW)
fivepairs<-c(15,37,46,151,181)
# Table 2.2 in Design of Observational Studies (DOS)
NSW[is.element(NSW$id,fivepairs),]
# Pair differences in earnings after treatment
dif<-NSW$re78[NSW$z==1]-NSW$re78[NSW$z==0]
# Chapter 2, footnote 7 of
stats::wilcox.test(dif,conf.int=TRUE)
```

---

periodontal   *Smoking and Periodontal Disease*

Description

Data from NHANES 2011-2012 containing 441 matched pairs of a daily cigarette smoker and a never smoker, recording the extent of periodontal disease. See Rosenbaum (2017) and Chapter 20 of "Design of Observational Studies", second edition.

Usage

```
data("periodontal")
```

Format

A data frame with 882 observations on the following 12 variables.

SEQN NHANES 2011-2012 sequence number

female =1 for female, 0 for male

age Age in years

black =1 for black, 0 for other

educf Education, in five categories. An ordered factor with levels <9 for less than 9th grade, 9 to 11 for 9th to 11th grade, HS/GED for high school or GED degree, SomeCol for some college, College for college degree.

income Ratio of family income to the poverty level, capped at 5 times the poverty level.

cigsperday Cigarettes smoked per day for daily smokers, 0 for never smokers

either Number of periodonal measurements indicative of periodontal disease.

neither Number of periodonal measurements

pcteither Percent indicative of periodontal disease, =100*either/neither.

z Treatment indicator, 1=daily smoker, 0=never smoker

mset Matched set indicator, 1 to 441.

Details

Excluding wisdom teeth, 6 measurements are taken for each tooth that is present, up to 28 teeth. Following Tomar and Asma (2000), a measurement indicates periodontal disease if either there is a loss of attachment of at least 4mm or a pocket depth of at least 4mm. The first individual has 11 measurements indicative of periodontal disease, out of 106 measurements, so pcteither is 100*11/106 = 10.38 percent. A related data set in DOS2 with bivariate outcome is teeth.

Source

Data are from the National Health and Nutrition Examination Survey 2011-2012 and were used as an example in Rosenbaum (2017). In the second edition of Design of Observational Studies, these data are discussed in Chapter 20, Evidence Factors.

References

Rosenbaum, P. R. (2015) <https://obsstudies.org/two-r-packages-for-sensitivity-analysis-in-observational-studies/> "Two R packages for sensitivity analysis in observational studies". Observational Studies, 1(1), 1-17.

Rosenbaum, P. R. (2017) <doi:10.1214/17-STS621> "The general structure of evidence factors in observational studies". Statistical Science 32, 514-530.

Tomar, S. L. and Asma, S. (2000) <doi:10.1902/jop.2000.71.5.743> "Smoking attributable periodontitis in the US: Findings from NHANES III". J Periodont 71, 743-751.

"US National Health and Nutrition Examination Survey 2011-2012". www.cdc.gov/nchs/nhanes/index.htm

Examples

```
# Figure 1 in Rosenbaum (2017)
data(periodontal)
attach(periodontal)
oldpar<-par()
m<-matrix(1:2,1,2)
layout(m,widths=c(1,2))
boxplot(pcteither[z==1]-pcteither[z==0],ylab="Smoker-Control Difference",
        main="(i)",xlab="Matched Pairs",ylim=c(-100,100))
abline(h=0,lty=2)
crosscutplot(cigsperday[z==1],pcteither[z==1]-pcteither[z==0],ylab="Smoker-Control Difference",
     xlab="Cigarettes per Day",main="(ii)",ylim=c(-100,100))
abline(h=0,lty=2)

# Sensitivity analysis in Section 2.3 of Rosenbaum (2017)
y<-pcteither[z==1]-pcteither[z==0]
x<-cigsperday[z==1]
senWilcox(y,gamma=2.76)
# The following is the same as sensitivitymw::senmw(y,gamma=2.77,method="p")
sensitivitymult::senm(pcteither,z,mset,gamma=2.77,inner=.5,trim=2)
# The following is the same as sensitivitymw::senmw(y,gamma=3.5,method="p")
sensitivitymult::senm(pcteither,z,mset,gamma=3.5,inner=.5,trim=2)
# Second evidence factor
crosscut(x,y)
crosscut(x,y,gamma=1.6)

# Note, however, that other statistics report greater insensitivity to
# bias by virtue of having larger design sensitivity:
sensitivitymult::senm(pcteither,z,mset,gamma=3.5,inner=1,trim=4)
sensitivitymult::senm(pcteither,z,mset,gamma=4.2,inner=1,trim=4)
senU(y,m1=4,m2=5,m=5,gamma=2.77)
senU(y,m1=6,m2=8,m=8,gamma=2.77)
senU(y,m1=6,m2=8,m=8,gamma=3.5)
detach(periodontal)
par(oldpar)
```

pinto	*Welding and DNA-Protein Crosslinks*

Description

This data set is from Pinto et al. (2000) and it describes 22 professional painters and 22 controls matched for age. All are men. The outcome is a measure of genetic damage, namely the frequency of micronuclei in 3000 oral epithelial cells scraped from the cheek, recorded as micronuclei per 1000 cells. The data are used as an example in Chapter 5 and 18 of "Design of Observational Studies", where the data illustrate the subtle relationship between doses of treatment and sensitivity to unmeasured biases. The dose of treatment is years of work as a professional painter, from 1.6 years to 40 years for painters, and it is highly correlated with age for painters, but it is zero for controls of all ages.

Usage

```
data("pinto")
```

Format

A data frame with 44 observations on the following 6 variables.

id ID number from Pinto et al. (2000).

pair Pair number, 1 to 22. The pair numbers are different from "Design of Observational Studies", where pairs were sorted by doses.

group painter or control

age age in years

years years of work as a professional painter, 0 for controls

longEx Painter in this pair had long exposure, TRUE if years>=4 years.

mn Micronuclei per 1000 cells

Source

The data are from Pinto et al. (2000). The data are used as an example in Chapter 5 of Design of Observational Studies. Chapter 18 shows that the pattern seen in this example is expected in theory, namely focusing on high-dose pairs makes the study more insensitive to unmeasured biases, despite the loss of sample size.

References

Pinto, D., J. M. Ceballos, G. Garcia, P. Guzman, L. M. Del Razo, E. Vera, H. Gomez, A. Garcia, and M. E. Gonsebatt (2000) <doi:10.1016/S1383-5718(00)00024-3> "Increased cytogenetic damage in outdoor painters". Mutation Research/Genetic Toxicology and Environmental Mutagenesis 467, 105-111.

Rosenbaum, P. R. (2003) <doi:10.1093/biostatistics/4.1.1> "Does a dose response relationship reduce sensitivity to hidden bias?". Biostatistics, 4, 1-10.

Rosenbaum, P. R. (2004) <doi:10.1093/biomet/91.1.153> "Design sensitivity in observational stud-
ies". Biometrika, 91, 153-164. Does design sensitivity calculations with doses of treatment.

Examples

```
data(pinto)
oldpar<-par(mfrow=c(1,3))
attach(pinto)
boxplot(mn~group,ylim=c(0,6),main="All",ylab="Micronuclei")
boxplot(mn[!longEx]~group[!longEx],ylim=c(0,6),main="Short Ex",ylab="Micronuclei")
boxplot(mn[longEx]~group[longEx],ylim=c(0,6),main="Long Ex",ylab="Micronuclei")

# Calculations in Table 5.5 of Design of Observational Studies (2010)
d<-mn[group=="painter"]-mn[group=="control"] # 22 pair differences
senWilcox(d,gamma=1)
senWilcox(d,gamma=2) # sensitive to gamma=2
senWilcox(d,gamma=3.3)
dLong<-d[longEx[group=="painter"]] # 12 pairs with long exposure
senWilcox(dLong,gamma=3.3) # insensitive to gamma=3.3
par(oldpar)
```

---

planScheffe	*Combining One Planned Comparison and a Scheffe Correction For All Comparisons.*

Description

The function planScheffe() computes the critical values for a level alpha test that combines one
planned linear combination of a K-dimensional multivariate Normal outcome and consideration of
all possible combinations correcting for multiple testing using a Scheffe projection. The function is
the same as the planScheffe() function in the sensitivitymult package, but the examples are different.
The method is discussed in Section 18.3 of "Design of Observational Studies", second edition, and
in Rosenbaum (2019).

Usage

```
planScheffe(K, alpha = 0.05)
```

Arguments

K	An integer >=2 giving the number of outcomes to be compared.
alpha	The level of the test, with 0 < alpha < 1.

Details

This method is discussed in section 18.3 of the second edition of "Design of Observational Studies".

Although the calculation uses the multivariate Normal distribution, a typical application uses K test
statistics that are asymptotically Normal.

The method is based on Rosenbaum (2019). The example below reproduces some of the compar-
isons in that manuscript.

Value

critical   critical is a vector with two elements, a and c. The null hypothesis is rejected
                at level alpha if either the Normal deviate for the planned comparison is >= a
                or if the square of the Normal deviate for any comparison is >= b. Then the
                probability of a false rejection is <= alpha.

alpha   alpha is a vector with three elements, a, c and joint. The value of joint should
                equal the input value of alpha aside from numerical errors of computation: it is
                the probability of a false rejection using the joint test that rejects if either of the
                two critical values in critical is exceeded. In contrast, a is the probability that
                the planned deviate will be >= critical[1] when the null hypothesis is true. Also,
                c is the probability that at least one comparison will have a squared deviate >=
                critical[2] when the null hypothesis is true.

Note

The method is based on Rosenbaum (2019).

The functions "cohere" may be used to calculate the standardized deviates that are compared to the
critical values from 'planScheffe'. The function "cohere" has options for an a priori comparison
or consideration of all possible comparisons with a Scheffe correction. The function "planScheffe"
provides a third option: one planned comparison plus all possible comparisons.

See also 'planScheffe' in the 'sensitivitymult' package.

Author(s)

Paul R. Rosenbaum.

References

Rosenbaum, P. R. (2016) <doi:10.1214/16-AOAS942> "Using Scheffe projections for multiple out-
comes in an observational study of smoking and periondontal disease". Annals of Applied Statistics,
10, 1447-1471.

Rosenbaum, P. R. (2019) <doi:10.1093/biostatistics/kxy055> "Combining planned and discovered
comparisons in observational studies". Biostatistics, to appear.

Scheffe, H. (1953) <doi:10.1093/biomet/40.1-2.87> "A method for judging all contrasts in the anal-
ysis of variance". Biometrika, 40, 87-104.

Examples

```
data(teeth)
attach(teeth)

planScheffe(2,alpha=0.05)

# Planned comparison w=c(1,1)
cohere(cbind(either4up,either4low),smoker,mset,
            w=c(1,1),gamma=2,m=8,m1=6,m2=8)

# Discovered comparison emphasizing upper teeth
```

```
cohere(cbind(either4up,either4low),smoker,mset,
        w=c(1,3),gamma=2,m=8,m1=6,m2=8)
3.465038^2 #squared deviate

# Both deviates lead to rejection, because
# 3.291909 >= 1.894915
# and 3.465038^2 = 12.00649 >= 7.077349

detach(teeth)
```

---

schoket   *DNA Damage in Aluminum Production Workers*

Description

This data set is from Schoket et al. (1991) and is discussed in Chapter 17 of "Design of Observational Studies", second edition. The data describe 25 aluminum production workers (w) and 25 controls (c) matched for age and smoking. The outcome is a measure of genetic damage, namely DNA adducts per 10^8 nucleotides. The data are used as an example in Chapter 17 of "Design of Observational Studies", where the data illustrate the possibility that some treated individuals are strongly affected by treatment, while others are unaffected, so the average treatment effect may be small, but the effect on affected individuals may be large and evident in data.

Usage

```
data("schoket")
```

Format

A data frame with 25 observations on the following 9 variables.

pair Pair number, 1 to 25.

idw Worker ID from Schoket et al. (1991).

agew Worker age in years

smokingw Worker cigarettes per day

adductsw Worker DNA adducts

idc Control ID from Schoket et al. (1991).

agec Control age in years

smokingc Control cigarettes per day

adductsc Control DNA adducts

Source

The data are from Schoket et al. (1991). The data are used as an example in Chapter 17 of "Design of Observational Studies"", second edition.

References

Conover, W. J. and Salsburg, D. S. (1988) <doi:10.2307/2531906> "Locally most powerful tests for detecting treatment effects when only a subset of patients can be expected to respond to treatment". Biometrics, 189-196.

Rosenbaum, P. R. (2007) <doi:10.1111/j.1541-0420.2007.00783.x> "Confidence intervals for uncommon but dramatic responses to treatment". Biometrics, 63(4), 1164-1171.

Rosenbaum, P. R. (2011) <doi:10.1111/j.1541-0420.2010.01535.x> "A new U statistic with superior design sensitivity in matched observational studies". Biometrics, 67(3), 1017-1027.

Schoket, B., Phillips, D. H., Hewer, A. and Vincze, I. (1991) <doi:10.1016/0165-1218(91)90084-Y> "32P-postlabelling detection of aromatic DNA adducts in peripheral blood lymphocytes from aluminium production plant workers". Mutation Research/Genetic Toxicology, 260(1), 89-98.

Stephenson, W. R. (1981) <doi:10.1080/01621459.1981.10477749> "A general class of one-sample nonparametric test statistics based on subsamples". Journal of the American Statistical Association, 76(376), 960-966.

Examples

```
data(schoket)
attach(schoket)
plot(sort(adductsc),sort(adductsw),ylim=c(0,6.4),xlim=c(0,6.4),
    xlab="DNA adducts for controls",ylab="DNA adducts for workers",
    main="Quantile-Quantile Plot") # Compare with Chapter 17
abline(0,1) # line of equality
legend(4,1,lty=1,"x=y")
boxplot(adductsw,adductsc,ylim=c(0,6.4),ylab="DNA adducts",names=c("Worker","Control"))
d<-adductsw-adductsc
senWilcox(d,gamma=1)
senWilcox(d,gamma=1.5) # sensitive to gamma=1.5
senU(d,gamma=1) # U-statistic version of Wilcoxon's statistic
senU(d,gamma=1.8)
# Stephenson's statistic is obtained from senU()
# by setting m1=m2=m
senU(d,m1=5,m2=5,m=5,gamma=1) # Stephenson's statistic, m=5
senU(d,m1=5,m2=5,m=5,gamma=1.8)
# U-statistic from Rosenbaum (2011) and
# Section 19.2 of Design of Observational Studies, 2nd ed.
senU(d,m1=4,m2=5,m=5,gamma=1.8)
detach(schoket)
```

---

senU   *Sensitivity Analysis for a New U Statistic*

Description

Sensitivity analysis for the new U statistic of Rosenbaum (2011). For m=m1=m2, this is the test of Stephenson (1981). The ranks proposed by Stephenson closely approximate the optimal ranks proposed by Conover and Salzburg (1988) for detecting a treatment that has a large effect on a

small subpopulation and no effect on most of the population; see Rosenbaum (2007). The example reproduces some results from Chapter 17 of "Design of Observational Studies", second edition, and the method is discusssed in Section 19.2.

Usage

```
senU(d, gamma = 1, m = 2, m1 = 2, m2 = 2, conf.int = FALSE,
     alpha = 0.05, alternative = "greater", exact = NULL)
```

Arguments

d	A vector of treated-minus-control matched pair differences in outcomes.
gamma	gamma >= 1 is the value of the sensitivity parameter.
m	See m2.
m1	See m2.
m2	If (m,m1,m2) are three integers such that 1 <= m1 <= m2 <= m, then the triple (m,m1,m2) defines a U statistic. If (m,m1,m2) = (1,1,1), then the U statistic is the sign test statistic. If (m,m1,m2) = (2,2,2), then it is the U statistic that closely approximates Wilcoxons signed rank test. If m=m1=m2, then the U statistic is the test of Stephenson (1981). The general U statistic is discussed in Rosenbaum (2011).
conf.int	If conf.int=TRUE, the a 1-alpha confidence interval and an interval of point estimates is returned in addition to the P-value testing no treatment effect.
alpha	Coverage rate of the confidence interval. With probability at least 1-alpha, the confidence interval will cover the treatment effect providing the bias in treatment assignment is at most gamma.
alternative	If alternative = "greater" or alternative = "less", then one-sided tests and intervals are returned. If alternative = "twosided", then both one sided tests are done, with the smaller P-value doubled to yield a two-sided P-value. If alternative = "twosided", the confidence interval is the intersection of two one-sided 1-alpha/2 confidence intervals.
exact	If exact is NULL, then exact is set to TRUE if length(d) <= 50, and is set to FALSE if length(d) > 50. The ranks used by the U statistic involve combinatorial coefficiencts that grow rapidly with increasing sample size. If exact=TRUE, these ranks are computed exactly using expression (8) in Rosenbaum (2011). If exact=FALSE, the ranks are computed by an asymptotic approximation that does not involve large combinatorial coefficients, specifically expression (9) in Rosenbaum (2012).

Details

This method is discussed in Chapter 17 and Section 19.2 of the second edition of "Design of Observational Studies".

The general method is developed in Rosenbaum (2011).

Value

`pval`	The upper bound on the P-value testing no effect in the presence of a bias in treatment assignment of at most gamma. If the bias in treatment assignment is at most gamma, and if there is no treatment effect, then there is at most an alpah chance that the P-value is less than alpha, this being true for all 0<alpha<1.
`estimate`	If conf.int=TRUE, the interval of point estimates of an additive treatment effect in the presence of a bias in treatment assigment of at most gamma. If gamma=1, then you are assuming ignorable treatment assignment or equivalently no unmeasured confounding, so the interval collapses to a point, and that point is the usual Hodges-Lehmann point estimate.
`conf.int`	If conf.int=TRUE, the a 1-alpha confidence interval for an additive treatment effect in the presence of a bias in treatment assignment of at most gamma. If gamma=1, then this is the usual confidence interval obtained by inverting the Wilcoxon test, and it would be appropriate in a paired randomized experiment.

Note

The test of Stephenson (1981) uses ranks similar to those of Conover and Salzburg (1988) which were designed to have high power when most people are unaffected by treatment, but a small subpopulation is strongly affected; see Rosenbaum (2007). This is the situation discussed in Chapter 17 of "Design of Observational Studies", second edition; see also Section 19.2. Even for pair differences d that are Normal with expectation tau and constant variance, the Wilcoxon test tends to exaggerate the degree of sensitivity to unmeasured bias. Compare the Wilcoxon test and the U statistic with (m,m1,m2) = (5,4,5) in the Normal situation. In a randomized experiment (gamma=1), the two tests have the same Pitman efficiency. However, as the number of pairs increases with tau=0.5, the Wilcoxon test has limiting sensitivity to bias of gamma=3.2 while (m,m1,m2) = (5,4,5) has limiting sensitivity 3.9, and (m,m1,m2) = (8,7,8) has limiting sensitivity 5.1. See Rosenbaum (2011) for specifics.

Author(s)

Paul R. Rosenbaum

References

Conover, W. J. and Salsburg, D. S. (1988) <doi:10.2307/2531906> "Locally most powerful tests for detecting treatment effects when only a subset of patients can be expected to" respond" to treatment". Biometrics, 189-196.

Hodges Jr, J. L. and Lehmann, E. L. (1963) <doi:10.1214/aoms/1177704172> "Estimates of location based on rank tests". The Annals of Mathematical Statistics, 598-611.

Rosenbaum, P. R. (1993) <doi:10.1080/01621459.1993.10476405> "Hodges-Lehmann point estimates of treatment effect in observational studies". Journal of the American Statistical Association, 88(424), 1250-1253.

Rosenbaum, P. R. (2007) <doi:10.1111/j.1541-0420.2007.00783.x> "Confidence intervals for uncommon but dramatic responses to treatment". Biometrics, 63(4), 1164-1171.

Rosenbaum, P. R. (2011) <doi:10.1111/j.1541-0420.2010.01535.x> "A new U statistic with superior design sensitivity in matched observational studies". Biometrics, 67(3), 1017-1027.

Schoket, B., Phillips, D. H., Hewer, A. and Vincze, I. (1991) <doi:10.1016/0165-1218(91)90084-Y> "32P-postlabelling detection of aromatic DNA adducts in peripheral blood lymphocytes from aluminium production plant workers". Mutation Research/Genetic Toxicology, 260(1), 89-98.

Stephenson, W. R. (1981) <doi:10.1080/01621459.1981.10477749> "A general class of one-sample nonparametric test statistics based on subsamples". Journal of the American Statistical Association, 76(376), 960-966.

Examples

```
data("schoket")
d<-schoket$adductsw-schoket$adductsc

# With the defaults, m=2, m1=2, m2=2, the U-statistic is very
# similar to Wilcoxon's signed rank statistic
senWilcox(d,gamma=1)
senU(d,gamma=1)

# With m=1, m1=1, m2=1, the U-statistic is the sign test
senU(d,gamma=1,m=1,m1=1,m2=1)
prop.test(sum(d>0),length(d),p=.5,
      alternative="greater",correct=FALSE)$p.value

# With m=m1=m2, this is the test of Stephenson (1981) whose ranks are similar to
# those of Conover and Salzburg (1988); see Rosenbaum (2007).

# The calculations that follow reproduce the sensitivity analysis for the
# data of Schoket et al. (1991) in Chapter 17 of "Design of
# Observational Studies", second edition.
senU(d,gamma=1,m=2,m1=2,m2=2)
senU(d,gamma=1,m=5,m1=5,m2=5)

senU(d,gamma=1.5,m=2,m1=2,m2=2)
senU(d,gamma=1.5,m=5,m1=5,m2=5)

senU(d,gamma=1.8,m=2,m1=2,m2=2)
senU(d,gamma=1.8,m=5,m1=5,m2=5)

senU(d,gamma=2,m=2,m1=2,m2=2)
senU(d,gamma=2,m=5,m1=5,m2=5)

data(lead)
# Reproduces parts of Table 2 in Rosenbaum (2011)
senU(lead$dif,gamma=5.8,m=8,m1=5,m2=8)
senU(lead$dif,gamma=5,m=5,m1=4,m2=5)

# m=2, m1=2, m2=2 is the U-statistic that closely
# resembles Wilcoxon's signed rank test.  Note
# that the results are almost the same.
senWilcox(lead$dif,gamma=5) # In Table 2
senU(lead$dif,gamma=5,m=2,m1=2,m2=2)
```

---

senWilcox   *Sensitivity Analysis for Wilcoxon's Signed-rank Statistic*

Description

Sensitivity analysis for Wilcoxon's signed rank statistic in observational studies. Performs a sensitivity analysis for the P-value, the Hodges-Lehmann estimate of an additive treatment effect, and the confidence interval for the effect. The example is from Section 3.5 of "Design of Observational Studies", 2nd ed., and it also illustrates the amplification in Section 3.6 using the 'amplify' function from the 'sensitivitymult' package.

Usage

```
senWilcox(d, gamma = 1, conf.int = FALSE, alpha = 0.05, alternative = "greater")
```

Arguments

d	A vector of treated-minus-control matched pair differences in outcomes.
gamma	gamma >= 1 is the value of the sensitivity parameter.
conf.int	If conf.int=TRUE, the a 1-alpha confidence interval and an interval of point estimates is returned in addition to the P-value testing no treatment effect.
alpha	Coverage rate of the confidence interval. With probability at least 1-alpha, the confidence interval will cover the treatment effect providing the bias in treatment assignment is at most gamma.
alternative	If alternative = "greater" or alternative = "less", then one-sided tests and intervals are returned. If alternative = "twosided", then both one sided tests are done, with the smaller P-value doubled to yield a two-sided P-value. If alternative = "twosided", the confidence interval is the intersection of two one-sided 1-alpha/2 confidence intervals.

Details

The senWilcox function uses a large sample Normal approximation to the distribution of Wilcoxon's signed rank statistic. When gamma=1, it should agree with the wilcox.test() function in the stats package with exact=FALSE and correct=FALSE.

The example reproduces the example of the large-sample approximation in Section 3.5 of "Design of Observational Studies". Note that the confidence intervals in Table 3.3 of that book are exact, not approximate, so they are slightly different.

Value

pval	The upper bound on the P-value testing no effect in the presence of a bias in treatment assignment of at most gamma. If the bias in treatment assignment is at most gamma, and if there is no treatment effect, then there is at most an alpha chance that the P-value is less than alpha, this being true for all 0<alpha<1.

estimate   If conf.int=TRUE, the interval of point estimates of an additive treatment effect in the presence of a bias in treatment assigment of at most gamma. If gamma=1, then you are assuming ignorable treatment assignment or equivalently no unmeasured confounding, so the interval collapses to a point, and that point is the usual Hodges-Lehmann point estimate.

conf.int   If conf.int=TRUE, the a 1-alpha confidence interval for an additive treatment effect in the presence of a bias in treatment assignment of at most gamma. If gamma=1, then this is the usual confidence interval obtained by inverting the Wilcoxon test, and it would be appropriate in a paired randomized experiment.

Note

The Wilcoxon test is not the best test for use in sensitivity analyses – it tends to exaggerate how sensitive a study is to bias, saying it is more sensitive than it truly is. Other, less familiar, test statistics avoid this issue; see Rosenbaum (2010, 2011, 2013, 2014, 2015). Learn about this using a 'shinyapp', at rosenbap.shinyapps.io/learnsenShiny/

Author(s)

Paul R. Rosenbaum

References

Hodges Jr, J. L. and Lehmann, E. L. (1963) <doi:10.1214/aoms/1177704172> "Estimates of location based on rank tests". The Annals of Mathematical Statistics, 598-611.

Hollander, M., Wolfe, D. and Chicken, E. (2013) <doi:10.1002/9781119196037> "Nonparametric Statistical Methods". (3rd edition) New York: John Wiley.

Lehmann, E. L. (1975). "Nonparametrics" <https://www.springer.com/gp/book/9780387352121> San Francisco: Holden-Day. Reprinted by Prentice-Hall and Springer.

Rosenbaum, P. R. (1987) <doi:10.1093/biomet/74.1.13> "Sensitivity analysis for certain permutation inferences in matched observational studies". Biometrika, 74(1), 13-26.

Rosenbaum, P. R. (1993) <doi:10.1080/01621459.1993.10476405> "Hodges-Lehmann point estimates of treatment effect in observational studies". Journal of the American Statistical Association, 88(424), 1250-1253.

Rosenbaum, P. R. (2002) <doi:10.1007/978-1-4757-3692-2_3> "Observational Studies", Second edition. New York: Springer. Wilcoxon's test is discussed in Section 4.3.3.

Rosenbaum, P. R. (2007) <doi:10.1111/j.1541-0420.2006.00717.x> "Sensitivity Analysis for M Estimates, Tests, and Confidence Intervals in Matched Observational Studies". Biometrics, 63(2), 456-464. R-packages 'sensitivitymult', 'sensitivitymv', 'sensitivityfull', 'sensitivitymw'

Rosenbaum, P. R. (2010) <doi:10.1198/jasa.2010.tm09570> "Design sensitivity and efficiency in observational studies". Journal of the American Statistical Association, 105(490), 692-702.

Rosenbaum, P. R. (2011) <doi:10.1111/j.1541-0420.2010.01535.x> "A new U statistic with superior design sensitivity in matched observational studies". Biometrics, 67(3), 1017-1027.

Rosenbaum, P. R. (2013) <doi:10.1111/j.1541-0420.2012.01821.x> "Impact of multiple matched controls on design sensitivity in observational studies". Biometrics, 69(1), 118-127. R-packages 'sensitivitymv', 'sensitivitymult' and 'sensitivityfull'

Rosenbaum, P. R. (2014) <doi:10.1080/01621459.2013.879261> "Weighted M statistics with superior design sensitivity in matched observational studies with multiple controls". Journal of the American Statistical Association, 109(507), 1145-1158. R-package 'sensitivitymw'

Rosenbaum, P. R. (2015) <doi:10.1080/01621459.2014.960968> "Bahadur efficiency of sensitivity analyses in observational studies". Journal of the American Statistical Association, 110(509), 205-217.

Examples

```
data(werfel)
d<-werfel$serpc_p-werfel$cerpc_p

# Reproduces the approximate one-sided P-value computed
# in Section 3.5 of "Design of Observational Studies".
senWilcox(d,gamma=3)

# Amplification in Section 3.6
sensitivitymult::amplify(3,5)
sensitivitymult::amplify(3,c(5,5.8,7))

# Reproduces parts of Tables 4.3 and 4.4 in Rosenbaum (2002)
data(lead)
senWilcox(lead$dif,gamma=1,conf.int=TRUE,alternative="twosided")
senWilcox(lead$dif,gamma=2,conf.int=TRUE,alternative="twosided")

# Agrees with the usual Wilcoxon procedures when gamma=1.
senWilcox(d,gamma=1,conf.int=TRUE,alternative="twosided")
stats::wilcox.test(d,conf.int=TRUE,exact=FALSE,correct=FALSE)
```

---

senWilcoxExact   *Exact Sensitivity Analysis for Wilcoxon's Signed-rank Statistic*

Description

Exact sensitivity analysis for Wilcoxon's signed rank statistic in observational studies. Performs a sensitivity analysis for the one-sided P-value. The method can be used in small samples without ties; however, it is primarily of theoretical interest, as the large sample approximation in 'senWilcox' is fine for most samples of practical size.

Usage

```
senWilcoxExact(d, gamma = 1)
```

Arguments

d	A vector of treated-minus-control matched pair differences in outcomes. There must be no ties in |d| when computing the exact distribution. If ties are present, use 'senWilcox' instead.
gamma	gamma >= 1 is the value of the sensitivity parameter.

Details

The exact method is discussed in Section 3.12 of "Design of Observational Studies", second edition. Tables 3.2 and 3.3 of Section 3.5 use these exact calculations.

Value

The upper bound on the one-sided, upper-tailed P-value testing no treatment effect in the presence of a bias in treatment assignment of at most gamma.

Note

The 'senWilcox' function uses a large-sample approximation, adding confidence intervals and point estimates.

Author(s)

Paul R. Rosenbaum

References

Pagano, M. and Tritchler, D. (1983) <doi:10.1080/01621459.1983.10477990> "On obtaining permutation distributions in polynomial time". Journal of the American Statistical Association, 78, 435-440.

Rosenbaum, P. R. (1987) <doi:10.1093/biomet/74.1.13> "Sensitivity analysis for certain permutation inferences in matched observational studies". Biometrika, 74(1), 13-26.

Examples

```
data(werfel)
d<-werfel$serpc_p-werfel$cerpc_p

# Reproduces the exact one-sided P-value computed in Section 3.9 of
# "Design of Observational Studies".
senWilcoxExact(d,gamma=2)

# Agrees with the usual Wilcoxon procedures when gamma=1.
senWilcoxExact(d,gamma=1)
stats::wilcox.test(d,alternative="greater")

# Reproduces the one-sided confidence interval for gamma=3 in Table 3.3
# of "Design of Observational Studies".
senWilcoxExact(d-0.0935,gamma=3)
senWilcoxExact(d-0.0936,gamma=3)
```

---

smahal   *Robust Mahalanobis Distance Matrix for Optimal Matching*

Description

Computes a robust Mahalanobis distance matrix between treated individuals and potential controls. The usual Mahalnaobis distance may ignore a variable because it contains one extreme outlier, and it pays too much attention to rare binary variables, but smahal addresses both issues. See Section 9.3 of "Design of Observational Studies", second edition.

Usage

```
smahal(z, X)
```

Arguments

z   z is a vector that is 1 for a treated individual and 0 for a control.

X   A matrix of continuous or binary covariates. The number of rows of X must
                equal the length of z.

Details

The usual Mahalnaobis distance may ignore a variable because it contains one extreme outlier, and it pays too much attention to rare binary variables, but smahal addresses both issues.

To address outliers, each column of x is replaced by a column of ranks, with average ranks used for ties. This prevents one outlier from inflating the variance for a column.

Rare binary variables have very small variances, $p(1-p)$ for small p, so in the usual Mahalanobis distance, a mismatch for a rare binary variable is counted as very important. If you were matching for US states as binary variables for individual states, mismatching for California would not be very important, because $p(1-p)$ is not so small, but mismatching for Wyoming is very important because $p(1-p)$ is very small. To combat this, the variances of the ranked columns are rescaled so they are all the same, all equal to the variance of untied ranks. See Chapter 9 of Design of Observational Studies, second edition.

Value

The robust distance matrix has one row for each treated individual (z=1) and one column for each potential control (z=0). The row and column names of the distance matrix refer to the position in z, 1, 2, ..., length(z).

Note

The matching functions in the 'DOS2' package are aids to instruction or self-instruction while reading "Design of Observational Studies". As in the book, these functions break the task of matching into small steps so they are easy to understand in depth. In practice, matching entails a fair amount of book-keeping best done by a package that automates these tasks. Consider R packages

'optmatch', 'rcbalance', 'DiPs', 'designmatch' or 'bigmatch'. Section 14.10 of "Design of Observational Studies", second edition, discusses and compares several R packages for optimal matching. Ruoqi Yu's 'DiPs' package computes robust distances.

Author(s)

Paul R. Rosenbaum

Examples

```
data(costa)
z<-1*(costa$welder=="Y")
aa<-1*(costa$race=="A")
smoker=1*(costa$smoker=="Y")
age<-costa$age
x<-cbind(age,aa,smoker)
dmat<-smahal(z,x)
# Mahalanobis distances
round(dmat[,1:6],2)
# Compare with Table 9.6 in "Design of
# Observational Studies", second edition
# Impose propensity score calipers
prop<-glm(z~age+aa+smoker,family=binomial)$fitted.values # propensity score
# Mahalanobis distanced penalized for violations of a propensity score caliper.
# This version is used for numerical work.
dmat<-addcaliper(dmat,z,prop,caliper=.5)
round(dmat[,1:6],2)
# Compare with Table 9.6 in "Design of
# Observational Studies", second edition

## Not run:
# You must load 'optmatch' to produce the match.
# Find the minimum distance match within propensity score calipers.
optmatch::pairmatch(dmat,data=costa)

## End(Not run)
# Conceptual versions with infinite distances
# for violations of propensity caliper.
dmat[dmat>20]<-Inf
round(dmat[,1:6],2)
```

---

tannery   *DNA Damage Among Tannery Workers*

Description

Data are from Zhang et al. (2008, Table 2) who studied DNA damage among tannery workers often exposed to trivalent chromium. The outcome is the mean tail moment (mtm) of the comet assay, a standard measure of DNA damage, with higher values signifying greater damage. The study describes 90 males in 30 blocks of 3 individuals. There are three groups, each with 30 individuals.

The three groups had a simlar distribution of ages, and the blocks control for smoking as closely as possible. Group e1 consists of 30 exposed workers at the tannery who worked in the tannery department, where the highest exposures to trivalent chromium are expected. Group e2 consists of 30 workers at the tannery who worked in the finishing department, where exposure to trivalent chromium is expected to be much lower. Group c consists of 30 controls who did not work at the tannery. This example is discussed in Chapter 20 of "Design of Observational Studies", 2nd ed.

Usage

```
data("tannery")
```

Format

A data frame with 30 observations on the following 4 variables.

block Block indicator, 1 to 30.

e1mtm mtm for the tannery worker from the tannery department.

e2mtm mtm for the tannery worker from the finishing department.

cmtm mtm for the control who did not work in the tannery

Details

The comet assay is described by Collins (2004). It is thought to measure DNA strand breaks, producing an image that resembles the tail of a comet, a larger, longer tail suggesting more extensive strand breaks. This example was discussed in Rosenbaum (2011).

Source

Data from Zhang et al. (2008, Table 2).

References

Collins, A. R. (2004) <doi:10.1385/MB:26:3:249> "The comet assay for DNA damage and repair: principles, applications, and limitations". Molecular Biotechnology 26(3), 249-261.

Rosenbaum, P. R. (2011) <doi:10.1198/jasa.2011.tm10422> "Some approximate evidence factors in observational studies". Journal of the American Statistical Association, 106, 285-293.

Rosenbaum, P. R. (2013) <doi:10.1111/j.1541-0420.2012.01821.x> "Impact of multiple matched controls on design sensitivity in observational studies". Biometrics 69 118-127. (Introduces inner trimming.)

Rosenbaum, P. R. (2015) <https://obsstudies.org/two-r-packages-for-sensitivity-analysis-in-observational-studies/> "Two R packages for sensitivity analysis in observational studies". Observational Studies, 1(1), 1-17.

Zhang, M., Chen, Z., Chen, Q., Zou, H., Lou, J. He, J. (2008) <doi:10.1016/j.mrgentox.2008.04.011> "Investigating DNA damage in tannery workers occupationally exposed to trivalent chromium using comet assay". Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 654(1), 45-51.

Examples

```
data(tannery)
boxplot(tannery[,2:4],names=c("Tannery E1","Finishing E2",
    "Control C"),ylab="Mean Tail Moment")
oldpar<-par(mfrow=c(1,2))
boxplot(tannery[,2:3],names=c("E1","E2"),ylab="Mean Tail Moment",
    main="Tannery vs. Finishing",ylim=c(0,12))
boxplot(as.vector(unlist(tannery[,2:3])),tannery[,4],
    names=c("E1+E2","C"),ylab="Mean Tail Moment",
    main="Exposed vs. Control",ylim=c(0,12))

# Stratified Wilcoxon analysis from the chapter Evidence Factors
# of Design of Observational Studies, Second Edition
# Also reproduces the F1, F2 example in Rosenbaum (2011, sec 6).
y<-tannery[,2:4]
rkc<-t(apply(y,1,rank)) # Ranks for (E1,E2,C)
sum(rkc[,3]) # Stratified rank sum for C in (E1, E2, C)
(35-60)/sqrt(20)

y<-tannery[,2:3]
rkc<-t(apply(y,1,rank)) # Ranks for (E1,E2)
sum(rkc[,2]) # Stratified rank sum for E2 in (E1, E2)

# Reorganize y for input to 'separable1v' from 'sensitivitymult'
# 'separable1v' is one-sided, looking for a large rank sum

# Factor 1
y<-tannery[,4:2]*(-1)
rkc<-t(apply(y,1,rank)) # Ranks for -y for (C,E2,E1)
sensitivitymult::separable1v(rkc,gamma=1)
# Test for C in (E1, E2, C)
(85-60)/sqrt(20)
(35-60)/sqrt(20)
sensitivitymult::separable1v(rkc,gamma=6)
# Test for C in (E1, E2, C)
p1<-sensitivitymult::separable1v(rkc,gamma=7)$pval

#Factor 2
y<-tannery[,3:2]*(-1)
rkc<-t(apply(y,1,rank)) # Ranks for -y for (E2,E1)
sensitivitymult::separable1v(rkc,gamma=1)
# Test for E2 in (E2, E1)
# Combine P-values using Fisher's method
sensitivitymv::truncatedP(c(1.134237e-08,0.001743502),trunc=1)

# Larger gammas
sensitivitymult::separable1v(rkc,gamma=1.7)
p2<-sensitivitymult::separable1v(rkc,gamma=2)$pval
# Combine P-values using Fisher's method
c(p1,p2)
sensitivitymv::truncatedP(c(p1,p2),trunc=1)
```

```
# Nearly reproduces calculations from Section 6 of Rosenbaum (2011)
# However, in Rosenbaum (2011), the second factor
# uses a pooled scale factor, whereas senm does not,
# so the result is very slightly different.
attach(tannery)
mset<-rep(block,3)
zC<-c(rep(0,60),rep(1,30))
z12<-c(rep(1,30),rep(0,30),rep(NA,30))
y<-c(e1mtm,e2mtm,cmtm)
detach(tannery)
use<-!is.na(z12)
# Factor 1
sensitivitymult::senm(y,zC,mset,gamma=1,
    alternative="less",trim=1)
sensitivitymult::senm(y,zC,mset,gamma=11.7,
    alternative="less",trim=1)
# Factor 2
sensitivitymult::senm(y[use],z12[use],mset[use],
    gamma=2,alternative="greater",trim=1)

# Combine two evidence factors
p1<-sensitivitymult::senm(y,zC,mset,gamma=12,
     alternative="less",trim=1)$pval
p2<-sensitivitymult::senm(y[use],z12[use],mset[use],gamma=3,
     alternative="greater",trim=1)$pval
c(p1,p2)
sensitivitymv::truncatedP(c(p1,p2),trunc=1)
# Combine p-values using Fisher's method

# Other psi-functions often have higher design
# sensitivity; see Rosenbaum (2013)
par(oldpar)
```

---

teeth   *Smoking and Periodontal Disease.*

Description

Data from NHANES 2011-2012 concerning periodonal disease in 441 matched pairs of smokers and nonsmokers.

Usage

```
data("teeth")
```

Format

A data frame with 882 observations on the following 4 variables.

mset Matched pair indicator: 1, 2, ..., 441.

smoker Treatment indicator: 1 if current smoker, 0 if never smoker

either4up Measure of periodontal disease for upper teeth; see Details.

either4low Measure of periodontal disease for lower teeth; see Details.

cigsperday Cigarettes smoked per day. Zero for nonsmokers.

Details

Smoking is believed to cause periodontal disease; see Tomar and Asma (2000). Using more recent data from NHANES 2011-2012, the data describe 441 matched pairs of a daily smoker and a never smoker. Daily smokers smoked every day of the last 30 days. Never smokers smoked fewer than 100 cigarettes in their lives, do not smoke now, and had no tobacco use in the previous five days.

Pairs are matched for education, income, age, gender and black race.

Measurements were made for up to 28 teeth, 14 upper, 14 lower, excluding 4 wisdom teeth. Pocket depth and loss of attachment are two complementary measures of the degree to which the gums have separated from the teeth. Pocket depth and loss of attachment are measured at six locations on each tooth, providing the tooth is present. A measurement at a location was taken to exhibit disease if it had either a loss of attachement >=4mm or a pocked depth >=4mm, so each tooth contributes a score from 0 to 6. Upper and lower are the number of measurements exhibiting disease for upper and lower teeth.

This example is from Rosenbaum (2016) where more information may be found.

The data are closely related to the periodontal data set, where the outcome is univariate. The teeth data are mentioned in Section 18.9 of "Design of Observational Studies", second edition.

Source

"National Health and Nutrition Examination Survey" (NHANES), 2011-2012. cdc.gov/nchs/nhanes

References

Rosenbaum, P. R. (2016) <doi:10.1214/16-AOAS942> "Using Scheffe projections for multiple outcomes in an observational study of smoking and periondontal disease". Annals of Applied Statistics, 10, 1447-1471.

Tomar, S. L. and Asma, S. (2000) <doi:10.1902/jop.2000.71.5.743> "Smoking attributable periodontitis in the United States: Findings from NHANES III". J. Periodont. 71, 743-751.

Examples

```
data(teeth)
summary(teeth)
# See also the examples in the documentation for 'cohere'.
```

---

werfel   *Welding Fumes and DNA Damage*

Description

This data set from Werfel et al. (1998) describes 39 electric arc welders and 39 controls matched for age and smoking. All are men. The outcome is a measure of genetic damage; specifically, erpcp_p is a measure of DNA single strand breakage and DNA-protein cross-links using elution rates through polycarbonate filters with proteinase K. The data are used as an example in Chapter 3 of Design of Observational Studies.

Usage

```
data("werfel")
```

Format

A data frame with 39 observations on the following 6 variables.

cage Age in years of the control in a matched pair.

csmoke NS=nonsmoker, S=smoker for the control in a pair

cerpc_p erpcp_p for the control in a pair

sage Age in years of the welder in a matched pair.

ssmoke NS=nonsmoker, S=smoker for the welder in a pair

serpc_p erpcp_p for the welder in a pair

Source

The data are from Werfel et al. (1998). It is used as an example in Section 3.5 of "Design of Observational Studies". It is also discussed in Fogarty (2019) and Rosenbaum (2007, 2015).

References

Fogarty, C. B. (2019) <doi:10.1080/01621459.2019.1632072> "Studentized Sensitivity Analysis for the Sample Average Treatment Effect in Paired Observational Studies". Journal of the American Statistical Association, to appear.

Rosenbaum, P. R. (1987) <doi:10.1093/biomet/74.1.13> "Sensitivity analysis for certain permutation inferences in matched observational studies". Biometrika, 74(1), 13-26.

Rosenbaum, P. R. (2007) <doi:10.1111/j.1541-0420.2006.00717.x> "Sensitivity analysis for M estimates, tests, and confidence intervals in matched observational studies". Biometrics, 63(2), 456-464.

Rosenbaum, P. R. (2015) <https://obsstudies.org/two-r-packages-for-sensitivity-analysis-in-observational-studies/> "Two R packages for sensitivity analysis in observational studies". Observational Studies, 1(1), 1-17.

Werfel, U., Langen, V., Eickhoff, I., Schoonbrood, J., Vahrenholz, C., Brauksiepe, A., Popp, W. and Norpoth, K. (1998) <doi:10.1093/carcin/19.3.413> "Elevated DNA single-strand breakage frequencies in lymphocytes of welders exposed to chromium and nickel". Carcinogenesis, 19(3), 413-418.

Examples

```
data(werfel)
d<-werfel$serpc_p-werfel$cerpc_p

# Reproduces the approximate one-sided P-value computed in Section 3.5
senWilcox(d,gamma=3)

# Amplification in Section 3.6
sensitivitymult::amplify(3,5)
sensitivitymult::amplify(3,c(5,5.8,7))

# Agrees with the usual large sample Wilcoxon procedures when gamma=1.
senWilcox(d,gamma=1,conf.int=TRUE,alternative="twosided")
stats::wilcox.test(d,conf.int=TRUE,exact=FALSE,correct=FALSE)
```

Index