

Package ‘sensitivityfull’

September 22, 2016

Type Package

Title Sensitivity Analysis for Full Matching in Observational Studies

Version 1.0.0

Author Paul R. Rosenbaum

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

Description Sensitivity to unmeasured biases in an observational study that is a full match.

Imports stats

License GPL-2

LazyData TRUE

NeedsCompilation no

Repository CRAN

Date/Publication 2016-09-22 09:22:58

R topics documented:

mscoref	1
senfm	3
separable1f	7

Index	9
--------------	----------

mscoref	<i>Computes M-scores</i>
---------	--------------------------

Description

Of limited interest to most users, mscoref computes the scores that form the basis for the hypothesis test performed by senfm.

Usage

```
mscoref(yamat, treated1, inner = 0, trim = 3, qu = 0.5)
```

Arguments

ymat	If there are I matched sets and the largest matched set contains J individuals, then y is an I by J matrix with one row for each matched set. If matched set i contains one treated individual and k controls, where k is at least 1 and at most J-1, then y[i,1] is the treated individual's response, y[i,2],...,y[i,k+1] are the responses of the k controls, and y[i,k+2],...,y[i,J] are equal to NA. If matched set i contains one control and k>1 treated individuals, then y[i,1] is the control's response, y[i,2],...,y[i,k+1] are the responses of the k treated individuals, and y[i,k+2],...,y[i,J] are equal to NA.
treated1	The vector treated1 is a logical vector of length I, where treated1[i]=TRUE if there is one treated subject in matched set i and treated1[i]=FALSE if there is more than one treated subject in matched set i.
inner	inner and trim together define the ψ -function for the M-statistic. The default values yield a version of Huber's ψ -function, while setting inner = 0 and trim = Inf uses the mean within each matched set. The ψ -function is an odd function, so $\psi(w) = -\psi(-w)$. For $w \geq 0$, the ψ -function is $\psi(w) = 0$ for $0 \leq w \leq \text{inner}$, is $\psi(w) = \text{trim}$ for $w \geq \text{trim}$, and rises linearly from 0 to trim for $\text{inner} < w < \text{trim}$.
trim	inner and trim together define the ψ -function for the M-statistic. See inner.
qu	Before applying the ψ -function to treated-minus-control differences, the differences are scaled by dividing by the qu quantile of all within set absolute differences. Typically, qu = 1/2 for the median. The value of qu has no effect if trim=0 and inner=Inf. See Maritz (1979) for the paired case and Rosenbaum (2007) for matched sets.

Value

Returns a matrix with the same dimensions as ymat and the same pattern of NAs. The returned value in position [i,j] compares ymat[i,j] to the other observations in row i of ymat, scoring the differences using ψ -function, totalling them, and applying a weight. Matched sets of unequal size are weighted using weights that would be efficient in a randomization test under a simple model with additive set and treatment effects and errors with constant variance. See Rosenbaum (2007, section 4.2).

When a matched set contains one control and several treated subjects, this is reflected in the returned scores by a sign reversal.

Author(s)

Paul R. Rosenbaum

References

- Huber, P. (1981) Robust Statistics. New York: John Wiley.
- Maritz, J. S. (1979). A note on exact robust confidence intervals for location. Biometrika 66 163–166.
- Rosenbaum, P. R. (2007). Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. Biometrics 63 456–64. (R package sensitivitymv)

Rosenbaum, P. R. (2010). Design of Observational Studies. New York: Springer. Table 2.12, page 60, illustrates the calculations for the simple case of matched pairs.

Rosenbaum, P. R. (2013). Impact of multiple matched controls on design sensitivity in observational studies. Biometrics 69 118-127. (Introduces inner trimming to increase design sensitivity.)

Examples

```
# The artificial example that follows has I=9
# matched sets. The first 3 sets have one treated
# individual and two controls with treated subjects
# in column 1. The next 3 sets are
# matched pairs, with treated subjects in column 1.
# The next 3 sets have one control and two treated
# subjects, with the control in column 1. Simulated
# from a Normal distribution with an additive effect
# of tau=1.

y<-c(2.2, 1.4, 1.6, 2.4, 0.7, 1.3, 1.2, 0.6, 0.3,
0.5, -0.1, -1.3, -0.3, 0.1, 0.4, 3.0, 1.1, 1.4, -0.8,
0.1, 0.8, NA, NA, NA, 1.1, 0.5, 1.8)
y<-matrix(y,9,3)
treated1<-c(rep(TRUE,6),rep(FALSE,3))

mscoref(y,treated1) # Huber scores
mscoref(y,treated1,inner=0.5,trim=3) #inner trimmed scores
mscoref(y,treated1,qu=.9,trim=1) #trimming the outer 10 percent

# For an additional example, install and load package sensitivitymv
# The following example is a match with variable controls.
# Both mscorev (in sensitivitymv) and mscoref (in sensitivityfull)
# reproduce the example in Rosenbaum (2007, Table 3).
# data(tbmetaphase)
# mscorev(tbmetaphase,trim=1)
# mscoref(tbmetaphase,rep(TRUE,15),trim=1)
```

senfm

Sensitivity analysis for a full match in an observational study.

Description

In a full match, each matched set contains either one treated individual and one or more controls or one control and one or more treated individuals. Uses Huber's M-statistic as the basis for the test, for instance, a mean. Performs either a randomization test or an analysis of sensitivity to departures from random assignment.

Usage

```
senfm(y, treated1, gamma = 1, inner = 0, trim = 3, lambda = 1/2,
      tau = 0, alternative="greater")
```

Arguments

y	<p>If there are I matched sets and the largest matched set contains J individuals, then y is an I by J matrix with one row for each matched set. If matched set i contains one treated individual and k controls, where k is at least 1 and at most J-1, then $y[i,1]$ is the treated individual's response, $y[i,2], \dots, y[i,k+1]$ are the responses of the k controls, and $y[i,k+2], \dots, y[i,J]$ are equal to NA. If matched set i contains one control and $k > 1$ treated individuals, then $y[i,1]$ is the control's response, $y[i,2], \dots, y[i,k+1]$ are the responses of the k treated individuals, and $y[i,k+2], \dots, y[i,J]$ are equal to NA.</p> <p>Exception: If y is a vector, not a matrix, then every matched set is understood to be a matched pair, one treated subject and one control, and y is understood to be a vector of $I = \text{length}(y)$ matched pair differences, treated-minus-control. The same results would be produced by setting y equal to an I-by-2 matrix whose first column contains treated responses and whose second column contains control responses such that the difference between the first and second column equals the vector y. For matched pairs, the package <code>sensitivitymw</code> does more, but <code>sensitivityfull</code> can be used.</p>
treated1	<p>The vector <code>treated1</code> is a logical vector of length I, where <code>treated1[i]=TRUE</code> if there is one treated subject in matched set i and <code>treated1[i]=FALSE</code> if there is more than one treated subject in matched set i.</p>
gamma	<p><code>gamma</code> is the sensitivity parameter Γ, where $\Gamma \geq 1$. Setting $\Gamma = 1$ is equivalent to assuming ignorable treatment assignment given the matched sets, and it performs a within-set randomization test.</p>
inner	<p><code>inner</code> and <code>trim</code> together define the ψ-function for the M-statistic. The default values yield a version of Huber's ψ-function, while setting <code>inner = 0</code> and <code>trim = Inf</code> uses the mean within each matched set. The ψ-function is an odd function, so $\psi(w) = -\psi(-w)$. For $w \geq 0$, the ψ-function is $\psi(w) = 0$ for $0 \leq w \leq \text{inner}$, is $\psi(w) = \text{trim}$ for $w \geq \text{trim}$, and rises linearly from 0 to <code>trim</code> for <code>inner</code> < w < <code>trim</code>.</p> <p>An error will result unless $0 \leq \text{inner} \leq \text{trim}$.</p> <p>Taking <code>trim</code> < <code>Inf</code> limits the influence of outliers; see Huber (1981). Taking <code>inner</code> > 0 often increases design sensitivity; see Rosenbaum (2013).</p>
trim	<p><code>inner</code> and <code>trim</code> together define the ψ-function for the M-statistic. See <code>inner</code>.</p>
lambda	<p>Before applying the ψ-function to treated-minus-control differences, the differences are scaled by dividing by the <code>lambda</code> quantile of all within set absolute differences. Typically, <code>lambda = 1/2</code> for the median. The value of <code>lambda</code> has no effect if <code>trim=0</code> and <code>inner=Inf</code>. See Maritz (1979) for the paired case and Rosenbaum (2007) for matched sets.</p> <p>An error will result unless $0 < \text{lambda} < 1$.</p>
tau	<p>The null hypothesis asserts that the treatment has an additive effect, <code>tau</code>. By default, <code>tau=0</code>, so by default the null hypothesis is Fisher's sharp null hypothesis of no treatment effect.</p>
alternative	<p>If <code>alternative="greater"</code>, the null hypothesis of a treatment effect of <code>tau</code> is tested against the alternative of a treatment effect larger than <code>tau</code>. If <code>alternative="less"</code>, the null hypothesis of a treatment effect of <code>tau</code> is tested against the alternative of</p>

a treatment effect smaller than tau. In particular, alternative="less" is equivalent to: (i) alternative="greater", (ii) y replaced by -y, and (iii) tau replaced by -tau. See the note for discussion of two-sided sensitivity analyses.

Details

For the given Γ , senfm computes the upper bound on the 1-sided P-value testing the null hypothesis of an additive treatment effect tau against the alternative hypothesis of a treatment effect larger than tau. By default, senfm tests the null hypothesis of no treatment effect against the alternative of a positive treatment effect. The P-value is an approximate P-value based on a Normal approximation to the null distribution; see Rosenbaum (2007).

Matched sets of unequal size are weighted using weights that would be efficient in a randomization test under a simple model with additive set and treatment effects and errors with constant variance; see Rosenbaum (2007).

The upper bound on the P-value is based on the separable approximation described in Gastwirth, Krieger and Rosenbaum (2000); see also Rosenbaum (2007).

Value

pval	The upper bound on the 1-sided P-value.
deviate	The deviate that was compared to the Normal distribution to produce pval.
statistic	The value of the M-statistic.
expectation	The maximum expectation of the M-statistic for the given Γ .
variance	The maximum variance of the M-statistic among treatment assignments that achieve the maximum expectation. Part of the separable approximation.

Note

The function senfm performs 1-sided tests. One approach to a 2-sided, α -level test does both 1-sided tests at level $\alpha/2$, and rejects the null hypothesis if either 1-sided test rejects. Equivalently, a bound on the two sided P-value is the smaller of 1 and twice the smaller of the two 1-sided P-values. This approach views a 2-sided test as two 1-sided tests with a Bonferroni correction; see Cox (1977, Section 4.2). In all cases, this approach is valid large sample test: a true null hypothesis is falsely rejected with probability at most α if the bias in treatment assignment is at most Γ ; so, this procedure is entirely safe to use. For a randomization test, $\Gamma = 1$, this Bonferroni procedure is not typically conservative. For large Γ , this Bonferroni procedure tends to be somewhat conservative.

Related packages are sensitivitymv, sensitivitymw and sensitivity2x2xk. sensitivitymv is for matched sets with one treated subject and a variable number of controls. sensitivitymw is for matched sets with one treated subject and a fixed number of controls, including matched pairs. For their special cases, sensitivitymv and sensitivitymw contain additional features not available in sensitivityfull. sensitivitymw is faster and computes confidence intervals and point estimates. sensitivitymw also implements methods from Rosenbaum (2014). sensitivity2x2xk is for 2x2xk contingency tables, treatment x outcome x covariates; see Rosenbaum and Small (2016).

Author(s)

Paul R. Rosenbaum.

References

- Cox, D. R. (1977). The role of significance tests (with Discussion). *Scand. J. Statist.* 4, 49-70.
- Hansen, B. B. (2007). Optmatch. *R News* 7 18-24. (R package optmatch) (Optmatch can create an optimal full match.)
- Hansen, B. B. and Klopfer, S. O. (2006). Optimal full matching and related designs via network flows. *J. Comput. Graph. Statist.* 15 609-627. (R package optmatch)
- Huber, P. (1981) *Robust Statistics*. New York: John Wiley. (M-estimates based on M-statistics.)
- Maritz, J. S. (1979). A note on exact robust confidence intervals for location. *Biometrika* 66 163–166. (Introduces exact permutation tests based on M-statistics by redefining the scaling parameter.)
- Rosenbaum, P. R. (1991). A characterization of optimal designs for observational studies. *J. Roy. Statist. Soc. B* 53 597-610. (Introduces full matching.)
- Rosenbaum, P. R. (2007). Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. *Biometrics* 63 456-64. (R package sensitivitymv)
- Rosenbaum, P. R. (2013). Impact of multiple matched controls on design sensitivity in observational studies. *Biometrics* 69 118-127. (Introduces inner trimming.)
- Rosenbaum, P. R. (2014). Weighted M-statistics with superior design sensitivity in matched observational studies with multiple controls. *J. Am. Statist. Assoc.* 109 1145-1158. (R package sensitivitymw)
- Rosenbaum, P. R. and Small, D. S. (2016). An adaptive Mantel-Haenszel test for sensitivity analysis in observational studies. *Biometrics*, to appear.

Examples

```
# The artificial example that follows has I=9
# matched sets. The first 3 sets have one treated
# individual and two controls with treated subjects
# in column 1. The next 3 sets are
# matched pairs, with treated subjects in column 1.
# The next 3 sets have one control and two treated
# subjects, with the control in column 1. Simulated
# from a Normal distribution with an additive effect
# of tau=1.

y<-c(2.2, 1.4, 1.6, 2.4, 0.7, 1.3, 1.2, 0.6, 0.3,
0.5, -0.1, -1.3, -0.3, 0.1, 0.4, 3.0, 1.1, 1.4, -0.8,
0.1, 0.8, NA, NA, NA, 1.1, 0.5, 1.8)
y<-matrix(y,9,3)
treated1<-c(rep(TRUE,6),rep(FALSE,3))

#Randomization test of no effect, Huber scores:
senfm(y,treated1)

#Sensitivity analysis, Huber scores:
senfm(y,treated1,gamma=2)

#Randomization test of tau=1 vs tau>1
senfm(y,treated1,tau=1)
```

```

#Randomization test of tau=1 vs tau<1
senfm(y,treated1,tau=1,alternative="less")

#Same randomization test of tau=1 vs tau<1
senfm(-y,treated1,tau=-1)

#Sensitivity analysis testing tau=1 at gamma=2
senfm(y,treated1,tau=1,gamma=2,alternative="greater")
senfm(y,treated1,tau=1,gamma=2,alternative="less")

# For an additional example, install and load package sensitivitymv
# The following example is a match with variable controls.
# So this example has one treated subject per matched set.
# Both mscorev (in sensitivitymv) and mscoref (in sensitivityfull)
# reproduce parts of the example in Rosenbaum (2007, Section 4).
# data(tbmetaphase)
# senmv(tbmetaphase,gamma=2,trim=1)
# senfm(tbmetaphase,rep(TRUE,15),gamma=2,trim=1)
# senmv(tbmetaphase,gamma=2,trim=1,tau=0.94)
# senfm(tbmetaphase,rep(TRUE,15),gamma=2,trim=1,tau=.94)
# senmv(tbmetaphase,gamma=2,trim=1,tau=0.945)
# senfm(tbmetaphase,rep(TRUE,15),gamma=2,trim=1,tau=.945)
# mscoref(tbmetaphase,rep(TRUE,15),trim=1)

```

separable1f

Computes the separable approximation.

Description

Of limited interest to most users, `separable1f` is called by the main function, `senfm`, in computing the sensitivity analysis. `separable1f` is given scores produced by `mscoref` and computes the separable approximation to the upper bound on the P-value.

Usage

```
separable1f(yamat, gamma = 1)
```

Arguments

<code>yamat</code>	A matrix of scores produced by <code>mscoref</code> .
<code>gamma</code>	The sensitivity parameter $\Gamma \geq 1$.

Details

See Gastwirth, Krieger and Rosenbaum (2000) and Rosenbaum (2007, section 4) for discussion of the separable approximation.

Value

pval	The upper bound on the 1-sided P-value.
deviate	The deviate that was compared to the Normal distribution to produce pval.
statistic	The value of the score-statistic.
expectation	The maximum expectation of the score-statistic for the given Γ .
variance	The maximum variance of the score-statistic among treatment assignments that achieve the maximum expectation. Part of the separable approximation.

Author(s)

Paul R. Rosenbaum

References

Gastwirth, J. L., Krieger, A. M. and Rosenbaum, P. R. (2000). Asymptotic separability in sensitivity analysis. *J. Roy. Statist. Soc. B.* 62 545-555.

Rosenbaum, P. R. (2007). Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. *Biometrics* 63 456-64. (See section 4.) (R package `sensitivitymv`)

Examples

```
# The artificial example that follows has I=9
# matched sets. The first 3 sets have one treated
# individual and two controls with treated subjects
# in column 1. The next 3 sets are
# matched pairs, with treated subjects in column 1.
# The next 3 sets have one control and two treated
# subjects, with the control in column 1. Simulated
# from a Normal distribution with an additive effect
# of tau=1.

y<-c(2.2, 1.4, 1.6, 2.4, 0.7, 1.3, 1.2, 0.6, 0.3,
0.5, -0.1, -1.3, -0.3, 0.1, 0.4, 3.0, 1.1, 1.4, -0.8,
0.1, 0.8, NA, NA, NA, 1.1, 0.5, 1.8)
y<-matrix(y,9,3)
treated1<-c(rep(TRUE,6),rep(FALSE,3))

# The same calculation done in two equivalent ways.
separable1f(mscoref(y,treated1),gamma=2)
senfm(y,treated1,gamma=2)
```

Index

*Topic **htest**
 senfm, 3
*Topic **robust**
 senfm, 3

mscoref, 1

senfm, 3
separable1f, 7