

Package ‘senstrat’

July 23, 2025

Type Package

Title Sensitivity Analysis for Stratified Observational Studies

Version 1.0.3

Author Paul R. Rosenbaum

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

Description Sensitivity analysis in unmatched observational studies, with or without strata. The main functions are `sen2sample()` and `senstrat()`. See Rosenbaum, P. R. and Krieger, A. M. (1990), JASA, 85, 493-498, <doi:10.1080/01621459.1990.10476226> and Gastwirth, Krieger and Rosenbaum (2000), JRSS-B, 62, 545–555 <doi:10.1111/1467-9868.00249> .

License GPL-2

Imports stats, BiasedUrn, MASS

Suggests sensitivitymw

Encoding UTF-8

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2017-07-16 21:01:17 UTC

Contents

compute	2
ev	3
evall	4
hodgeslehmann	5
homocyst	7
mscores	8
sen2sample	10
senstrat	12
zeta	16
Index	18

compute _p	<i>Computes individual and pairwise treatment assignment probabilities.</i>
----------------------	---

Description

Of limited interest to most users, the compute_p function plays an internal role in 2-sample and stratified sensitivity analyses. The compute_p function is equations (9) and (10), page 496, in Rosenbaum and Krieger (1990).

Usage

```
computep(bigN, n, m, g)
```

Arguments

bigN	Total sample size in this stratum.
n	Treated sample size in this stratum.
m	The number of 1's in the vector <i>u</i> of unobserved covariates. Here, <i>u</i> has bigN-m 0's followed by m 1's.
g	The sensitivity parameter Γ , where $\Gamma \geq 1$.

Value

p1	Equation (9), page 496, in Rosenbaum and Krieger (1990) evaluated with $u[i]=1$.
p0	Equation (9), page 496, in Rosenbaum and Krieger (1990) evaluated with $u[i]=0$.
p11	Equation (10), page 496, in Rosenbaum and Krieger (1990) evaluated with $u[i]=1, u[j]=1$.
p10	Equation (10), page 496, in Rosenbaum and Krieger (1990) evaluated with $u[i]=1, u[j]=0$.
p00	Equation (10), page 496, in Rosenbaum and Krieger (1990) evaluated with $u[i]=0, u[j]=0$.

Note

The function compute_p is called by the function ev.

Author(s)

Paul R. Rosenbaum

References

Rosenbaum, P. R. and Krieger, A. M. (1990). Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*, 85, 493-498.

Rosenbaum, P. R. (2002). *Observational Studies* (2nd edition). New York: Springer. Section 4.6.

Examples

```
compute(10,5,6,2)
```

 ev

Computes the null expectation and variance for one stratum.

Description

Of limited interest to most users, the `ev` function plays an internal role in 2-sample and stratified sensitivity analyses. The expectation and variance returned by the `ev` function are defined in the third paragraph of section 4, page 495, of Rosenbaum and Krieger (1990).

Usage

```
ev(sc, z, m, g, method)
```

Arguments

<code>sc</code>	A vector of scored outcomes for one stratum. For instance, for Wilcoxon's rank sum test, these would be the ranks of the outcomes in the current stratum.
<code>z</code>	Treatment indicators, with $\text{length}(z) = \text{length}(sc)$. Here, $z[i] = 1$ if i is treated and $z[i] = 0$ if i is control.
<code>m</code>	The unobserved covariate u has $\text{length}(z) - m$ 0's followed by m 1's.
<code>g</code>	The sensitivity parameter Γ , where $\Gamma \geq 1$.
<code>method</code>	If <code>method="RK"</code> or if <code>method="BU"</code> , exact expectations and variances are used in a large sample approximation. Methods "RK" and "BU" should give the same answer, but "RK" uses formulas from Rosenbaum and Krieger (1990), while "BU" obtains exact moments for the extended hypergeometric distribution using the <code>BiasedUrn</code> package and then applies Proposition 20, page 155, section 4.7.4 of Rosenbaum (2002). In contrast, <code>method="LS"</code> does not use exact expectations and variances, but rather uses the large sample approximations in section 4.6.4 of Rosenbaum (2002). Finally, <code>method="AD"</code> uses <code>method="LS"</code> for large strata and <code>method="BU"</code> for smaller strata.

Details

The function `ev()` is called by the function `eval()`.

Value

<code>expect</code>	Null expectation of the test statistic.
<code>vari</code>	Null variance of the test statistic.

Author(s)

Paul R. Rosenbaum

References

Rosenbaum, P. R. and Krieger, A. M. (1990). Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*, 85, 493-498.

Rosenbaum, P. R. (2002). *Observational Studies* (2nd edition). New York: Springer. Section 4.6.

Examples

```
ev(1:5, c(0, 1, 0, 1, 0), 3, 2, "RK")
ev(1:5, c(0, 1, 0, 1, 0), 3, 2, "BU")
```

evall

Compute expectations and variances for one stratum.

Description

Of limited interest to most users, the `evall()` function plays an internal role in 2-sample and stratified sensitivity analyses. The expectation and variance returned by the `evall()` function are defined in the third paragraph of section 4, page 495, of Rosenbaum and Krieger (1990). The function `evall()` calls the function `ev()` to determine the expectation and variance of the test statistic for an unobserved covariate u with $\text{length}(z)-m$ 0's followed by m 1's, doing this for $m=1, \dots, \text{length}(z)-1$.

Usage

```
evall(sc, z, g, method)
```

Arguments

<code>sc</code>	A vector of scored outcomes for one stratum. For instance, for Wilcoxon's rank sum test, these would be the ranks of the outcomes in the current stratum.
<code>z</code>	Treatment indicators, with $\text{length}(z)=\text{length}(sc)$. Here, $z[i]=1$ if i is treated and $z[i]=0$ if i is control.
<code>g</code>	The sensitivity parameter Γ , where $\Gamma \geq 1$.
<code>method</code>	If <code>method="RK"</code> or if <code>method="BU"</code> , exact expectations and variances are used in a large sample approximation. Methods "RK" and "BU" should give the same answer, but "RK" uses formulas from Rosenbaum and Krieger (1990), while "BU" obtains exact moments for the extended hypergeometric distribution using the <code>BiasedUrn</code> package and then applies Proposition 20, page 155, section 4.7.4 of Rosenbaum (2002). In contrast, <code>method="LS"</code> does not use exact expectations and variances, but rather uses the large sample approximations in section 4.6.4 of Rosenbaum (2002). Finally, <code>method="AD"</code> uses <code>method="LS"</code> for large strata and <code>method="BU"</code> for smaller strata.

Details

The `evall()` function is called by the `sen2sample()` function and the `senstrat()` function.

Value

A data.frame with $\text{length}(z)-1$ rows and three columns. The first column, *m*, gives the number of 1's in the unobserved covariate vector, *u*. The second column, *expect*, and the third column, *var*, give the expectation and variance of the test statistic for this *u*.

Note

The example is from Table 1, page 497, of Rosenbaum and Krieger (1990). The example is also Table 4.15, page 146, in Rosenbaum (2002). The example refers to Cu cells. The data are originally from Skerfving et al. (1974).

Author(s)

Paul R. Rosenbaum

References

Rosenbaum, P. R. and Krieger, A. M. (1990). Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*, 85, 493-498.

Rosenbaum, P. R. (2002). *Observational Studies* (2nd edition). New York: Springer. Section 4.6.

Examples

```
z<-c(rep(0,16),rep(1,23))
CuCells<-c(2.7, .5, 0, 0, 5, 0, 0, 1.3, 0, 1.8, 0, 0, 1.0, 1.8,
           0, 3.1, .7, 4.6, 0, 1.7, 5.2, 0, 5, 9.5, 2, 3, 1, 3.5,
           2, 5, 5.5, 2, 3, 4, 0, 2, 2.2, 0, 2)
evall(rank(CuCells),z,2,"RK")
```

hodgeslehmann

Computes Hodges-Lehmann Aligned Ranks.

Description

Computes Hodges-Lehmann (1962) aligned ranks for use in stratified sensitivity analyses. For instance, the scores, *sc*, used in function *senstrat()* might be Hodges-Lehmann aligned ranks.

Usage

```
hodgeslehmann(y,z,st,align="median",tau=0)
```

Arguments

<code>y</code>	A vector outcome. An error will result if <code>y</code> contains NAs.
<code>z</code>	Treatment indicators, with $\text{length}(z)=\text{length}(y)$. Here, $z[i]=1$ if i is treated and $z[i]=0$ if i is control.
<code>st</code>	A vector of stratum indicators with $\text{length}(st)=\text{length}(y)$. The vector <code>st</code> may be numeric, such as 1,2,..., or it may be a factor.
<code>align</code>	Within each stratum, the observations are aligned or centered by subtracting a location measure computed from observations in that stratum. If <code>align="median"</code> , then the median is subtracted. If <code>align="mean"</code> , then the mean is subtracted. If <code>align="huber"</code> , then the Huber's m-estimate is subtracted, as computed by the function <code>huber()</code> in the MASS package. If <code>align="hl"</code> , then the one-sample Hodges-Lehmann estimate is subtracted, as suggested by Mehrotra et al. (2010), and as computed by the <code>wilcox.test()</code> function in the stats package. The <code>wilcox.test()</code> command may generate numerous warnings about situations that are not hazardous, such as ties. When <code>align="hl"</code> , warnings produced by <code>wilcox.test()</code> are suppressed.
<code>tau</code>	If <code>tau=0</code> , then the null hypothesis is Fisher's sharp null hypothesis of no treatment effect. If <code>tau</code> is nonzero, then <code>tau</code> is subtracted from the <code>y</code> 's for treated responses before aligning and ranking. If <code>tau</code> is nonzero, the null hypothesis is that the treatment has a constant additive effect <code>tau</code> , the same for all strata.

Value

A vector of $\text{length}(y)$ containing the aligned ranks.

Author(s)

Paul R. Rosenbaum

References

- Hodges, J. L. and Lehmann, E. L. (1962) Rank methods for combination of independent experiments in analysis of variance. *Annals of Mathematical Statistics*, 33, 482-497.
- Lehmann, E. L. (1975) *Nonparametrics*. San Francisco: Holden-Day.
- Mehrotra, D. V., Lu, X., and Xiaoming, L. (2010). Rank-based analysis of stratified experiments: alternatives to the van Elteren test. *American Statistician*, 64, 121-130.

Examples

```
data("homocyst")
attach(homocyst)
sc<-hodgeslehmann(log2(homocysteine),z,stf,align="hl")
summary(sc)
length(sc)
detach(homocyst)
```

homocyst

*Homocysteine levels in daily smokers and never smokers.***Description**

Data from NHANES 2005-2006 concerning homocysteine levels in daily smokers ($z=1$) and never smokers ($z=0$), aged 20 and older. Daily smokers smoked every day for the last 30 days, smoking an average of at least 10 cigarettes per day. Never smokers smoked fewer than 100 cigarettes in their lives, do not smoke now, and had no tobacco use in the previous 5 days.

Usage

```
data("homocyst")
```

Format

A data frame with 2475 observations on the following 10 variables.

SEQN 2005-2006 NHANES ID number.

homocysteine Homocysteine level, umol/L. Based on LBXHCY.

z $z=1$ for a daily smoker, $z=0$ for a never smoker. Based on SMQ020, SMQ040, SMD641, SMD650, SMQ680.

stf A factor for strata indicating female, age, education, BMI and poverty.

st Numeric strata indicating female, age, education, BMI and poverty.

female 1=female, 0=male. Based on RIAGENDR

age3 Three age categories, 20-39, 40-50, ≥ 60 . Based on RIDAGEYR.

ed3 Three education categories, <High School, High School, at least some College. Based on DMDDEDUC2.

bmi3 Three of the body-mass-index, BMI, <30, [30,35), ≥ 35 . Based on BMXBMI.

pov2 TRUE=income at least twice the poverty level, FALSE otherwise.

Details

Bazzano et al. (2003) noted higher homocysteine levels in smokers than in nonsmokers. See also Pimentel et al. (2016) for a related analysis. The example below is from Rosenbaum (2017).

Source

NHANES, the US National Health and Nutrition Examination Survey, 2005-2006.

References

- Bazzano, L. A., He, J., Muntner, P., Vupputuri, S. and Whelton, P. K. (2003) 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) Constructed second control groups and attenuation of unmeasured biases. *Journal of the American Statistical Association*, 111, 1157-1167.
- Rosenbaum, P. R. (2017) Sensitivity analysis for stratified comparisons in an observational study of the effect of smoking on homocysteine levels. Manuscript.

Examples

```
data(homocyst)
#Homocysteine levels for daily smokers and nonsmokers.
boxplot(log(homocyst$homocysteine)~homocyst$z)
```

mscores	<i>Computes M-Scores for Two-Sample or Stratified Permutation Inference.</i>
---------	--

Description

Computes M-scores for two-sample or stratified permutation inference and sensitivity analyses. For instance, the scores may be used in functions `sen2sample()` and `senstrat()`.

Usage

```
mscores(y, z, st = NULL, inner = 0, trim = 3, lambda = 0.5,
        tau = 0)
```

Arguments

- | | |
|-------|---|
| y | A vector outcome. An error will result if y contains NAs. |
| z | Treatment indicators, with $\text{length}(z)=\text{length}(y)$. Here, $z[i]=1$ if i is treated and $z[i]=0$ if i is control. |
| st | For an unstratified two-sample comparison, $st = \text{NULL}$. For a stratified comparison, st is a vector of stratum indicators with $\text{length}(st)=\text{length}(y)$. The vector st may be numeric, such as 1,2,..., or it may be a factor. |
| inner | See trim below. |
| trim | The two parameters, $\text{inner} \geq 0$ and $\text{trim} > \text{inner}$, determine the odd ψ -function of the M-statistic. Huber favored an M-statistic similar to a trimmed mean, with $\text{inner}=0$, $\text{trim}>0$, so that $\psi(y)$ is $\max(-\text{trim}, \min(\text{trim}, y))$. Setting $\text{inner}>0$ ignores y near zero and is most useful for matched pairs; see Rosenbaum (2013). In general, $\psi(y)$ is $\text{sign}(y) * \text{trim} * \min(\text{trim} - \text{inner}, \max(0, \text{abs}(y) - \text{inner})) / (\text{trim} - \text{inner})$, so it is zero for y in the interval $[-\text{inner}, \text{inner}]$, and rises linearly from 0 to trim on the interval $y = \text{inner}$ to $y = \text{trim}$, and equals trim for $y > \text{trim}$. If $\text{trim} = \text{Inf}$, then no trimming is done, and $\psi(y) = y$. |

lambda	The ψ -function is applied to a treated-minus-control difference in responses after scaling by the lambda quantile of the within-strata absolute differences. Typically, lambda=1/2 for the median.
tau	If tau=0, then the null hypothesis is Fisher's sharp null hypothesis of no treatment effect. If tau is nonzero, then tau is subtracted from the y's for treated responses before scoring the y's. If tau is nonzero, the null hypothesis is that the treatment has a constant additive effect tau, the same for all strata.

Value

A vector of length(y) containing the M-scores. The M-scores may be used in `senstrat()` or `sen2Sample()`.

Note

The i th response in stratum s , $R[si]$, is compared to the j th response in stratum s , $R[sj]$, as $\psi((R[si]-R[sj])/\sigma)$, and for each fixed i these values are averaged over the $n[s]-1$ choices of j in stratum s , where $n[s]$ is the size of stratum s , thereby producing the M-score for $R[si]$. Here, σ is the lambda quantile, usually the median, of $|R[si]-R[sj]|$, taken over all within stratum comparisons. This extension to stratified comparisons of the method of Maritz (1979) for matched pairs is described in Rosenbaum (2007, 2017).

Author(s)

Paul R. Rosenbaum

References

- Huber, P. (1981) Robust Statistics. New York: Wiley, 1981. M-statistics were developed by Huber.
- Maritz, J. S. (1979) Exact robust confidence intervals for location. *Biometrika* 1979, 66, 163-166. Maritz proposed small adjustments to M-statistics for matched pairs that permit them to be used in exact permutation tests and confidence intervals.
- Rosenbaum, P. R. (2007) Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. *Biometrics*, 2007, 63, 456-464. <doi:10.1111/j.1541-0420.2006.00717.x> Extends the method of Maritz (1979) to matching with multiple controls and to sensitivity analysis in observational studies.
- Rosenbaum, P. R. (2013) Impact of multiple matched controls on design sensitivity in observational studies. *Biometrics*, 2013, 69, 118-127. Computes the design sensitivity of M-statistics and proposes the trim parameter for use with matched pairs to increase design sensitivity.
- Rosenbaum, P. R. (2014) Weighted M-statistics with superior design sensitivity in matched observational studies with multiple controls. *Journal of the American Statistical Association* 109, 1145-1158. Discusses weights for matched sets to increase design sensitivity.
- Rosenbaum, P. R. (2017) Sensitivity analysis for stratified comparisons in an observational study of the effect of smoking on homocysteine levels. Manuscript. Among other things, extends the method of Maritz (1979) to stratified comparisons.

See Also

The packages `sensitivitymv`, `sensitivitymw` and `sensitivityfull` use M-scores in matched sets. The M-scores from those packages are similar, but are weighted differently, particularly when matched sets have varying sizes.

Examples

```
data("homocyst")
attach(homocyst)
sc<-mscores(log2(homocysteine),z,st=stf)
par(mfrow=c(1,2))
boxplot(log2(homocysteine)~z,main="Data")
boxplot(sc~z,main="Mscores")
detach(homocyst)
```

sen2sample

Treatment versus control sensitivity analysis without strata.

Description

Performs a two-sample, treated-versus-control sensitivity analysis without strata or matching. The method is described in Rosenbaum and Krieger (1990) and Rosenbaum (2002, Section 4.6). The example in those references is used below to illustrate use of the `sen2sample()` function.

Usage

```
sen2sample(sc, z, gamma = 1, alternative = "greater", method="BU")
```

Arguments

<code>sc</code>	A vector of scored outcomes for one stratum. For instance, for Wilcoxon's rank sum test, these would be the ranks of the outcomes.
<code>z</code>	Treatment indicators, with $\text{length}(z)=\text{length}(sc)$. Here, $z[i]=1$ if i is treated and $z[i]=0$ if i is control.
<code>gamma</code>	The sensitivity parameter Γ , where $\Gamma \geq 1$.
<code>alternative</code>	If <code>alternative="greater"</code> , then the test rejects for large values of the test statistic. If <code>alternative="less"</code> then the test rejects for small values of the test statistic. In a sensitivity analysis, it is safe but somewhat conservative to perform a two-sided test at level α by doing two one-sided tests each at level $\alpha/2$.
<code>method</code>	If <code>method="RK"</code> or if <code>method="BU"</code> , exact expectations and variances are used in a large sample approximation. Methods "RK" and "BU" should give the same answer, but "RK" uses formulas from Rosenbaum and Krieger (1990), while "BU" obtains exact moments for the extended hypergeometric distribution using the <code>BiasedUrn</code> package and then applies Proposition 20, page 155, section 4.7.4 of Rosenbaum (2002). In contrast, <code>method="LS"</code> does not use exact expectations and variances, but rather uses the large sample approximations in section 4.6.4 of Rosenbaum (2002). Finally, <code>method="AD"</code> uses <code>method="LS"</code> for large strata and <code>method="BU"</code> for smaller strata.

Value

sc	A vector of scored outcomes for one stratum. For instance, for Wilcoxon's rank sum test, these would be the ranks of the outcomes in the current stratum.
z	Treatment indicators, with $\text{length}(z)=\text{length}(sc)$. Here, $z[i]=1$ if i is treated and $z[i]=0$ if i is control.
g	The sensitivity parameter Γ , where $\Gamma \geq 1$.

Note

The example is from Table 1, page 497, of Rosenbaum and Krieger (1990). The example is also Table 4.15, page 146, in Rosenbaum (2002). The data are originally from Skerfving et al. (1974).

Author(s)

Paul R. Rosenbaum

References

- Rosenbaum, P. R. and Krieger, A. M. (1990). Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*, 85, 493-498.
- Rosenbaum, P. R. (2002). *Observational Studies* (2nd edition). New York: Springer. Section 4.6.
- Skerfving, S., Hansson, K., Mangs, C., Lindsten, J., & Ryman, N. (1974). Methylmercury-induced chromosome damage in man. *Environmental Research*, 7, 83-98.

See Also

For binary responses, use the `sensitivity2x2xk` package. For matched responses, use one of the following packages: `sensitivitymult`, `sensitivitymv`, `sensitivitymw`, `sensitivityfull`.

Examples

```
mercury<-c(5.3, 15, 11, 5.8, 17, 7, 8.5, 9.4, 7.8, 12, 8.7, 4, 3, 12.2, 6.1, 10.2,
          100, 70, 196, 69, 370, 270, 150, 60, 330, 1100, 40, 100, 70,
          150, 200, 304, 236, 178, 41, 120, 330, 62, 12.8)
z<-c(rep(0,16),rep(1,23))
CuCells<-c(2.7, .5, 0, 0, 5, 0, 0, 1.3, 0, 1.8, 0, 0, 1.0, 1.8,
          0, 3.1, .7, 4.6, 0, 1.7, 5.2, 0, 5, 9.5, 2, 3, 1, 3.5,
          2, 5, 5.5, 2, 3, 4, 0, 2, 2.2, 0, 2)

#Reproduces Rosenbaum and Krieger (1990), page 497
sen2sample(rank(mercury),z,gamma=5)
#Reproduces Rosenbaum and Krieger (1990), page 497
sen2sample(rank(CuCells),z,gamma=2)
(551.500000-492.334479)/sqrt(1153.775252) #Computation of the deviate
#Intermediate calculations: expectation and variance are in row 21.
eval1(rank(CuCells),z,2,method="RK")

#The following three examples, if run, reproduce the
#calculations in the final paragraph of page 145,
```

```
#Section 4.6.6 of Rosenbaum (2002) Observational Studies, 2nd Ed.
#The first calculation uses large sample approximations
#to expectations and variances.
sen2sample(rank(mercury),z,gamma=2,method="LS")
#The next two calculations use exact expectations and variances
sen2sample(rank(mercury),z,gamma=2,method="RK")
sen2sample(rank(mercury),z,gamma=2,method="BU")
```

senstrat

Sensitivity Analysis With Strata.

Description

Performs a sensitivity analysis with strata. The underlying method is described in Rosenbaum (2017), and the main example illustrates calculations from that paper. Use `sen2sample()` if there are no strata.

Usage

```
senstrat(sc, z, st, gamma = 1, alternative = "greater",
         level = 0.05, method="BU", detail = FALSE)
```

Arguments

sc	A vector of scored outcomes. For instance, these scored outcomes might be produced by the <code>hodgeslehmann()</code> function or the <code>mscores()</code> function. An error will result if <code>sc</code> contains NAs.
z	Treatment indicators, with <code>length(z)=length(sc)</code> . Here, <code>z[i]=1</code> if <code>i</code> is treated and <code>z[i]=0</code> if <code>i</code> is control.
st	Vector of stratum indicators, with <code>length(st)=length(sc)</code> . The vector <code>st</code> may be numeric, say 1, 2, ..., or it may be a factor. A factor will be converted to integers for computations. If there is only one stratum with data, then it is better to use <code>sen2sample</code> , not <code>senstrat</code> , and a warning will be given.
gamma	The sensitivity parameter Γ , where $\Gamma \geq 1$.
alternative	If <code>alternative="greater"</code> , then the test rejects for large values of the test statistic. If <code>alternative="less"</code> then the test rejects for small values of the test statistic. In a sensitivity analysis, it is safe but somewhat conservative to perform a two-sided test at level α by doing two one-sided tests each at level $\alpha/2$.
level	The sensitivity analysis is for a test of the null hypothesis of no treatment effect performed with the stated level, conventionally <code>level=0.05</code> . If there is no treatment effect, so the null hypothesis is true, and if the bias in treatment assignment is at most Γ , then the chance that the sensitivity analysis will falsely reject the null hypothesis is at most 0.05 if <code>level=0.05</code> . It would be common to report that rejection of the null hypothesis at conventional <code>level=0.05</code> is insensitive to a bias of Γ if this Γ is the largest Γ leading to rejection. Determining this largest Γ would entail running <code>senstrat</code> several times with different values of Γ . It is

	perfectly reasonable, if less conventional, to conduct a sensitivity analysis with some other level, say $\text{level}=0.01$, and this might be necessary if one were testing several hypotheses, correcting for multiple testing using the Bonferroni or Holm procedures.
method	If <code>method="RK"</code> or if <code>method="BU"</code> , exact expectations and variances are used in a large sample approximation. Methods "RK" and "BU" should give the same answer, but "RK" uses formulas from Rosenbaum and Krieger (1990), while "BU" obtains exact moments for the extended hypergeometric distribution using the BiasedUrn package and then applies Proposition 20, page 155, section 4.7.4 of Rosenbaum (2002). In contrast, <code>method="LS"</code> does not use exact expectations and variances, but rather uses the large sample approximations in section 4.6.4 of Rosenbaum (2002). Finally, <code>method="AD"</code> uses <code>method="LS"</code> for large strata and <code>method="BU"</code> for smaller strata.
detail	If <code>detail=FALSE</code> , concise practical output is produced. The option <code>detail=TRUE</code> provides additional details about the computations, which may be useful in understanding the computations or in trouble shooting, but the additional details are not useful in data analysis.

Details

The method uses a Normal approximation to the distribution of the test statistic. If method is not "LS", then this approximation is suitable for either a few strata containing many people or for many strata each stratum containing only a few people. In contrast, `method="LS"` is useful only if every single stratum contains a large sample.

Value

Conclusion	An English sentence stating the conclusion of the sensitivity analysis. This sentence says whether the null hypothesis has been rejected at the stated level, defaulting to $\text{level}=0.05$, in the presence of a bias of at most Γ . This statement is correct, if perhaps ever so slightly conservative. This statement should be understood as the conclusion produced by <code>senstrat</code> , with the remaining output being descriptive or approximate.
Result	Numeric results, including an approximate P-value, the deviate that was compared with the Normal distribution to produce the P-value, the test statistic formed as the sum of the scores for treated individuals, its null expectation and variance, and the value of Γ .
Description	The number of nondegenerate strata used in computations and the total number of treated individuals and controls in those strata.
StrataUse	An English sentence stating whether all strata were used or alternatively describing degenerate strata that were not used. See the Note.
LinearBound	If <code>detail=TRUE</code> , then the Result above is labeled as the LinearBound; it is a safe but perhaps slightly conservative P-value.
Separable	If <code>detail=TRUE</code> , then separable approximation of Gastwirth et al. (2000) is reported; it is a slightly liberal P-value. In many if not most examples, LinearBound and Separable are in close agreement, so the issue of liberal versus conservative does not arise. The stated Conclusion is based on the conservative LinearBound.

Remark	If detail=TRUE, then Remark contains an English sentence commenting upon the agreement of the LinearBound and the Separable approximation.
lambda	If detail=TRUE, then the values of the separable $\lambda(b)$ and the linear Taylor bound $\lambda(b) + \sum \eta_s$ from Rosenbaum (2017). The Remark above says there is agreement if these two quantities have the same sign.

Note

Strata that contain only treated subjects or only controls do not affect permutation inferences. These strata are removed before computations begin. Inclusion of these strata would not alter the permutation inference. A message will indicate whether any strata have been removed; see StrataUse in the value section above. You can avoid strata that do not contribute by using full matching in place of conventional stratification; see Rosenbaum (1991) and Hansen (2004) and R packages optmatch and sensitivityfull.

The output produces a rejection or acceptance of the null hypothesis at a stated level= α in the presence of a bias of at most Γ . This statement is entirely safe, in the sense that it is at worst a tad conservative, falsely rejecting a true null hypothesis with probability at most α in the presence of a bias of at most Γ . To produce a true P-value, you would need to run the program several times to find the smallest level= α that leads to rejection, and the P-value produced in this standard way would share the property of the test in being, at worst, slightly conservative. To save time, the output contains an approximate P-value that agrees with the accept/reject decision, but if this P-value is much smaller than the level – say, rejection at 0.05 with a P-value of 0.00048, then unlike the reject/accept decision, the 0.00048 P-value may not be conservative. I have never found the approximate P-value to be misleading. However, having seen an approximate P-value of 0.00048, it is easy to check whether you are formally entitled to reject at $\alpha=0.0005$ by rerunning the program with level=0.0005 and basing the conclusion on the reject/accept decision at level $\alpha=0.0005$.

When there are many small strata, Gastwirth, Krieger and Rosenbaum (2000, GKR) proposed a separable approximation to the sensitivity bound. In principle, this separable approximation is a tad liberal: it does not find the absolute worst unobserved covariate u , but rather a very bad u , such that as the number of strata increases the difference between the worst u and the very bad u becomes negligible. The current function senstrat() improves upon the separable approximation in the following way. This improvement is discussed in Rosenbaum (2017). It makes a one-step linear Taylor correction to the separable approximation which is guaranteed to be slightly conservative, rather than slightly liberal, so it is always safe to use: it falsely rejects at level = 0.05 with probability at most 0.05 in the presence of a bias of at most Γ . More precisely, unlike the method of GKR, the one-step LinearBound correction does not require many small strata: in large samples, it falsely rejects at level=0.05 with probability at most 0.05 whether there are few or many strata, even if some of the strata are much larger than others. If detail=FALSE, conclusions are based on the LinearBound without further comments. This is reasonable, because the LinearBound is safe to use in all cases, being at worst slightly conservative. If detail=TRUE, the LinearBound and the Separable approximation are compared. Usually, the LinearBound and the Separable approximation yield conclusions that are very close, providing some reassurance that the LinearBound is not very conservative and the Separable approximation is not very liberal. The option detail=TRUE is an aid to someone who wants to understand the LinearBound, but it is not a tool required for data analysis.

Author(s)

Paul R. Rosenbaum

References

- Gastwirth, J. L., Krieger, A. M., and Rosenbaum, P. R. (2000) Asymptotic separability in sensitivity analysis. *Journal of the Royal Statistical Society B* 2000, 62, 545-556. <doi:10.1111/1467-9868.00249>
- Hansen, B. B. (2004) Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association*, 99, 609-618. Application of full matching as an alternative to conventional stratification. See also Hansen's R package `optmatch`.
- Rosenbaum, P. R. and Krieger, A. M. (1990). Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*, 85, 493-498.
- Rosenbaum, P. R. (1991) A characterization of optimal designs for observational studies. *Journal of the Royal Statistical Society B*, 53, 597-610. Introduces full matching as an alternative to conventional stratification.
- Rosenbaum, P. R. (2002). *Observational Studies* (2nd edition). New York: Springer. Section 4.6.
- Rosenbaum, P. R. (2007) Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. *Biometrics*, 2007, 63, 456-464. <doi:10.1111/j.1541-0420.2006.00717.x> See the `erpcp` example below.
- Rosenbaum, P. R. (2014) Weighted M-statistics with superior design sensitivity in matched observational studies with multiple controls. *Journal of the American Statistical Association*, 109, 1145-1158. <doi:10.1080/01621459.2013.879261> Contains the mercury example below, 397 matched triples.
- Rosenbaum, P. R. and Small, D. S. (2017) An adaptive Mantel-Haenszel test for sensitivity analysis in observational studies. *Biometrics*, 73, 422-430. <doi:10.1111/biom.12591> The 2x2x2 BRCA example from Satagopan et al. (2001) in this paper can be used to compare the `senstrat()` function and the `mhLS()` function in the `sensitivity2x2xk` package for the Mantel-Haenszel test. The 2x2x2 table is in the documentation for `mhLS()`, but must be reformatted as individual data for use by `senstrat`. With binary outcomes, the extreme unobserved covariate is known from theory. Not knowing this theory, `senstrat()` computes the same answer as `mhLS()` for $\gamma=7$.
- Rosenbaum, P. R. (2017) Sensitivity analysis for stratified comparisons in an observational study of the effect of smoking on homocysteine levels. Manuscript.
- Satagopan, J. M., Offit, K., Foulkes, W., Robson, M. E. Wacholder, S., Eng, C. M., Karp, S. E. and Begg, C. B. (2001). The lifetime risks of breast cancer in Ashkenazi Jewish carriers of *brca1* and *brca2* mutations. *Cancer Epidemiology, Biomarkers and Prevention*, 10, 467-473.
- Werfel, U., Langen, V., Eickhoff, I. et al. Elevated DNA strand breakage frequencies in lymphocytes of welders exposed to chromium and nickel. *Carcinogenesis*, 1998, 19, 413-418. Used in the `erpcp` example below.

See Also

If outcomes are binary, then use the `sensitivity2x2xk` package. If there are no strata – that is, if everyone is in the same stratum, so there is a single stratum – then use the function `sen2sample()` in this package. If the strata are matched pairs or matched sets, then use one of the packages `sensitivitymult`, `sensitivityfull`, `sensitivitymv`, `sensitivitymw`.

Examples

```
data("homocyst")
attach(homocyst)
sc<-hodgeslehmann(log2(homocysteine),z,stf,align="hl")
senstrat(sc,z,stf,gamma=1.8)
# Compare this with:
senstrat(sc,z,stf,gamma=1.8,detail=TRUE)
# With detail=TRUE, it is seen that the separable and Taylor bounds
# on the maximum P-value are nearly identical. The Taylor upper bound
# is safe -- i.e., at worst conservative -- in all cases.
detach(homocyst)
#
# The following example compares senmw in the sensitivitymw package
# to senstrat in an example with 397 matched triples, one treated,
# two controls. We expect the separable approximation to work well
# with S=397 small strata, and indeed the results are identical.
library(sensitivitymw)
data(mercury)
senmw(mercury,gamma=15)
# Reformat mercury for use by senstrat().
z<-c(rep(1,397),rep(0,397),rep(0,397))
st<-rep(1:397,3)
y<-as.vector(as.matrix(mercury))
sc<-mscores(y,z,st=st)
senstrat(sc,z,st,gamma=15,detail=TRUE)
# The separable approximation from senmw() and senstrat() are identical,
# as they should be, and the Taylor approximation in senstrat()
# makes no adjustment to the separable approximation.
#
# The following example from the sensitivitymw package
# is for 39 matched pairs, so the separable algorithm
# and the Taylor approximation are not needed, yet
# they both provide exactly the correct answer.
library(sensitivitymw)
data(erpcp)
senmw(erpcp,gamma=3)
# Reformat erpcp for use by senstrat().
z<-c(rep(1,39),rep(0,39))
st<-rep(1:39,2)
y<-as.vector(as.matrix(erpcp))
sc<-mscores(y,z,st=st)
senstrat(sc,z,st,gamma=3,detail=TRUE)
```

zeta

zeta function in sensitivity analysis

Description

Of limited interest to most users, the zeta function plays an internal role in 2-sample and stratified sensitivity analyses. The zeta function is equation (8), page 495, in Rosenbaum and Krieger (1990).

Usage

```
zeta(bigN, n, m, g)
```

Arguments

bigN	Total sample size in this stratum.
n	Treated sample size in this stratum.
m	The number of 1's in the vector u of unobserved covariates. Here, u has bigN-m 0's followed by m 1's.
g	The sensitivity parameter Γ , where $\Gamma \geq 1$.

Value

The value of the zeta function.

Note

The zeta function is called by computep.

Author(s)

Paul R. Rosenbaum

References

Rosenbaum, P. R. and Krieger, A. M. (1990). Sensitivity of two-sampler permutation inferences in observational studies. *Journal of the American Statistical Association*, 85, 493-498.

Rosenbaum, P. R. (2002). *Observational Studies* (2nd edition). New York: Springer. Section 4.6.

Examples

```
zeta(10,5,6,2)
```

Index

- * **Causal inference**
 - sen2sample, 10
 - senstrat, 12
- * **Hodges-Lehmann aligned ranks**
 - senstrat, 12
- * **M-estimate**
 - mscores, 8
 - senstrat, 12
- * **M-test**
 - mscores, 8
- * **Observational study**
 - sen2sample, 10
 - senstrat, 12
- * **Sensitivity analysis**
 - sen2sample, 10
 - senstrat, 12
- * **Stratification**
 - mscores, 8
 - senstrat, 12
- * **Two-sample sensitivity analysis**
 - sen2sample, 10
- * **datasets**
 - homocyst, 7
- * **htest**
 - ev, 3
 - evall, 4
 - hodgeslehmann, 5
 - mscores, 8
 - sen2sample, 10
 - senstrat, 12
- * **rbounds**
 - sen2sample, 10
 - senstrat, 12
- * **robust**
 - ev, 3
 - evall, 4
 - hodgeslehmann, 5
 - mscores, 8
 - sen2sample, 10
 - senstrat, 12
- compute, 2
- ev, 3
- evall, 4
- hodgeslehmann, 5
- homocyst, 7
- mscores, 8
- sen2sample, 10
- senstrat, 12
- zeta, 16