# Package 'sensitivitymv'

November 20, 2013

**Type** Package

**Title** Sensitivity Analysis in Observational Studies

**Version** 1.2

**Date** 2013-11-20

**Author** Paul R. Rosenbaum

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

**Description**

Sensitivity analysis in observational studies, including evidence factors and amplification, using the permutation distribution of Huber-Maritz M-statistics, including the permutational t-test.

**License** GPL-2

**LazyLoad** yes

## R topics documented:

`sensitivitymv-package`

*Sensitivity analysis for observational studies*

## Description

Sensitivity analysis for matched observational studies with one or more controls using Huber-Maritz M-tests or the permutational t-test.

## Details

| | |
|---|---|
| Package: | sensitivitymv |
| Type: | Package |
| Version: | 1.0 |
| Date: | 2013-11-03 |
| License: GPL-2 LazyLoad: | yes |

The main function in the sensitivitymv package is senmv.

The sensitivitymv package performs sensitivity analyses in matched observational studies using Huber-Maritz M-tests, including the permutational t-test. There can be matched treatment-control pairs, or matched sets with one treated subject and a fixed number of controls, or matched sets with one treated subject and a variable number >=1 of controls. The package name, sensitivitymv, refers to sensitivity analyses with M-statistics and variable controls.

The sensitivity analysis asks about the magnitude, gamma, of bias in treatment assignment in observational studies that would need to be present to alter the conclusions of a randomization test that assumed matching for observed covariates removes all bias. The method implemented in sensitivitymv is essentially the method described in Rosenbaum (2007); see also Rosenbaum (2013). For general discussion of senstivity analyses in observational studies, see Chapter 4 of Rosenbaum (2002).

M-tests perform hypothesis tests using the statistics that Huber equated to zero in defining his M-estimates. The mean is one of many possible M-statistics. As developed by Maritz (1979), M-tests differ in some details from M-estimates. In particular, Maritz uses a scale factor that is fixed under the null hypothesis. Maritz developed his method for matched pairs, but it easily extends to matching with multiple or variable controls; see Rosenbaum (2007, Section 4).

The permutational t-test permutes the observations themselves rather than permuting robust scores derived from the observations. It is, of course, of historical interest, but it also continues to attract current attention because of its close connections to the average treatment effect and to the average effect of the treatment on the treated. See the method="t" option in senmv.

The main function in the sensitivitymv package is senmv. Other functions are truncatedP and truncatedPbv that combine independent P-values using the truncated product of Zaykin et al. (2002) and the amplify function that interprets values of gamma using the method in Rosenbaum and Silber (2009).

In a straightforward way, the senmv package may be used in calculations for approximate evidence factors in the sense of Rosenbaum (2011); see documentation for the truncatedP or truncatedPbv functions.

There are six data sets, erpcp, tbmetaphase, mercury, mtm, lead150 and lead250 . As noted in the documentation for senmv and truncatedP, these three data sets may be used to reproduce analyses

from the cited literature, as illustrated in the examples for senmv and truncatedP. The documentation for mscorev shows how to reproduce an intermediate result, specifically Table 3 in Rosenbaum (2007).

## Author(s)

Paul Rosenbaum

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

## References

Huber, P. (1981) Robust Statistics. New York: Wiley, 1981.

Maritz, J. S. (1979) Exact robust confidence intervals for location. Biometrika 1979, 66, 163-166.

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

Rosenbaum, P. R. (2007) 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) Amplification of sensitivity analysis in observational studies. Journal of the American Statistical Association, 104, 1398-1405.

Rosenbaum, P. R. (2011) Some approximate evidence factors in observational studies. Journal of the American Statistical Association, 106, 285-295.

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

Zaykin, D. V., Zhivotovsky, L. A., Westfall, P. H. and Weir, B. S. (2002) Truncated product method of combining P-values. Genetic Epidemiology, 22, 170-185.

---

`amplify` *Amplification of sensitivity analysis in observational studies.*

---

## Description

Uses the method in Rosenbaum and Silber (2009) to interpret a value of the sensitivity parameter gamma, for instance the parameter in the senmv function. Each value of gamma amplifies to a curve (lambda,delta) in a two-dimensional sensitivity analysis, the inference being the same for all points on the curve. That is, a one-dimensional sensitivity analysis in terms of gamma has a two-dimensional interpretation.

## Usage

```
amplify(gamma, lambda)
```

## Arguments

| | |
|---|---|
| `gamma` | gamma > 1 is the value of the sensitivity parameter, for instance the parameter in senmv. length(gamma)>1 will generate an error. |
| `lambda` | lambda is a vector of values > gamma. An error will result unless lambda[i] > gamma > 1 for every i. |

## Details

A single value of gamma, say gamma = 3.5 in the example, corresponds to a curve of values of (lambda, delta), including (4, 26), (6,8), (8,6), and (11,5) in the example. An unobserved covariate that is associated with a lambda = 6 fold increase in the odds of treatment and a delta = 8 fold increase in the odds of a positive pair difference is equivalent to gamma = 3.5.

The curve is gamma = (lambda*delta + 1)/(lambda+delta). Amplify is given one gamma and a vector of lambdas and solves for the vector of deltas. The calculation is elementary.

This interpretation of gamma is developed in detail in Rosenbaum and Silber (2009), and it makes use of Wolfe's (1974) family of semiparametric deformations of an arbitrary symmetric distribuiton.

Strictly speaking, the amplification describes matched pairs, not matched sets. The senmv function views a k-to-1 matched set with k controls matched to one treated individual as a collection of k correlated treated-minus-control matched pair differences; see Rosenbaum (2007). For matched sets, it is natural to think of the amplification as describing any one of the k matched pair differences in a k-to-1 matched set.

The curve has asymptotes that the function amplify does not compute: gamma corresponds with (lambda,delta) = (gamma, Inf) and (Inf, gamma).

A related though distict idea is developed in Gastwirth et al (1998). The two approaches agree when the outcome is binary, that is, for McNemar's test.

## Value

Returns a vector of values of delta of length(lambda) with names lambda.

## Note

The example expands the discussion of Table 1 in Rosenbaum (2007). The study is insensitive to a bias of gamma = 3.5. An unobserved covariate associated with a lambda = 6 fold increase in the odds of treatment and a delta= 8 fold increase in the odds of positive pair difference is equivalent to gamma = 3.5. Also, gamma = 3.5 is equivalent to (lambda,delta) = (4,26), (8,6) and (11,5).

## Author(s)

Paul R. Rosenbaum

## References

Gastwirth, J. L., Krieger, A. M., Rosenbaum, P. R. (1998) Dual and simultaneous sensitivity analysis for matched pairs. Biometrika, 85, 907-920.

Rosenbaum, P. R. (2007) 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) Amplification of sensitivity analysis in observational studies. Journal of the American Statistical Association, 104, 1398-1405.

Wolfe, D. A. (1974) A charaterization of population weighted symmetry and related results. Journal of the American Statistical Association, 69, 819-822.

## Examples

```
data(erpcp)
senmv(erpcp,gamma=3.5,trim=1)
amplify(3.5,6)
amplify(3.5,c(4,6,8,11))
```

## erpcp *DNA Damage Among Welders*

### Description

Matched pairs of a welder and a control, matching for age and smoking. The values are DNA elution rates through polycarbonate filters with proteinase K (or erpcp). Data are originally from Werfel et al. (1998) and were used as an example in Rosenbaum (2007). Data are used to illustrate the senmv function in the sensitivitymv package.

### Usage

```
data(erpcp)
```

### Format

A data frame with 39 observations on the following 2 variables.

`welder` erpcp value for the welder

`control` erpcp value for the matched control

### Source

Werfel et al. (1998).

### References

Rosenbaum, P. R. Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. Biometrics, 2007, 63, 456-464.

Werful, 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.

### Examples

```
data(erpcp)
```

## lead150 *Smoking and lead in 150 matched 1-5 sets.*

### Description

Comparison of 150 daily smokers and 5 matched control nonsmokers from NHANES in terms of blood lead levels in mu-g/L. The matching controlled for gender, age, race education level, and household income. The data were used in Rosenbaum (2013) to contrast design sensitivity with matched pairs and with multiple controls. See also lead250. Used as an example to illustrate the senmv function in the sensitivitymv package.

### Usage

```
data(lead150)
```

### Format

Column 1 is the lead level for the daily smoker. Columns 2 through 6 are lead levels for the five matched control nonsmokers.

### Details

See the appendix to Rosenbaum (2013).

### Source

NHANES 2007-2008 files from CDC and Rosenbaum (2013). Under a conventional Gaussian model, 250 pairs and 150 1-5 matched sets should have the same standard error; however, the design sensitivity is not the same.

### References

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

### Examples

```
data(lead150)
```

---

## lead250 *Smoking and lead in 250 matched pairs.*

---

### Description

Comparison of 250 daily smokers and one matched control nonsmoker from NHANES in terms of blood lead levels in mu-g/L. The matching controlled for gender, age, race education level, and household income. The data were used in Rosenbaum (2013) to contrast design sensitivity with matched pairs and with multiple controls. See also lead150. Used as an example to illustrate the senmv function in the sensitivitymv package.

### Usage

```
data(lead250)
```

### Format

Column 1 describes the smoker and column 2 describes the matched nonsmoker.

### Details

See the appendix to Rosenbaum (2013). Under a conventional Gaussian model, 250 pairs and 150 1-5 matched sets should have the same standard error; however, the design sensitivity is not the same.

### Source

NHANES 2007-2008 from CDC and Rosenbaum (2013).

## References

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

## Examples

```
data(lead250)
```

---

`mercury` *NHANES Mercury/Fish Data*

---

## Description

Data from NHANES. Matrix 397 x 3. n=397 treated people who ate at least 15 servings of fish or shellfish during the previous month are matched to two controls who ate at most one serving of fish or shellfish. The values in mercury record the level of methylmercury in blood in mu-g/dl. Column 1 is treated, columns 2 and 3 are controls. Specifically, column 2 describes an individual who ate no fish or shellfish in the previous month, while column 3 describes an individual who ate exactly one serving of fish or shellfish. Data are used to illustrate the senmv function in the sensitivitymv package.

## Usage

```
data(mercury)
```

## Format

A data frame with 397 observations on the following 3 variables.

`Treated` Mercury level for treated individual.

`Zero` Mercury level for a control who ate no fish/shellfish

`One` Mercury level for a control who ate one serving of fish/shellfish

## Source

Data is originally from NHANES 2009-2010.

## Examples

```
data(mercury)
```

`mscorev` *Computes the M-scores used by senmv.*

## Description

Computes the M-scores used by senmv for sensitivity analysis in observational studies using Huber-Maritz M-tests.

## Usage

```
mscorev(ymat, inner = 0, trim = 2.5, qu = 0.5, TonT = FALSE)
```

## Arguments

| | |
|---|---|
| `ymat` | ymat is a matrix as described in the documentation for senmv. |
| `inner` | inner is the parameter described in the documentation for senmv. |
| `trim` | trim is the parameter described in the documentation for senmv. |
| `qu` | qu is the lambda parameter described in the documentation for senmv. |
| `TonT` | If TonT=FALSE, then the total score in set (row) i is divided by the number ni of individuals in row i, as in expression (8) in Rosenbaum (2007). If TonT=TRUE, then the division is by ni-1, not by ni, and there is a further division by the total number of matched sets. See the discussion of TonT in the documentation for senmv. |

## Value

Generally, a matrix with the same dimensions as ymat containing the M-scores. Exception: if a matched set does not contain at least one treated subject and at least one control, then that set will not appear in the result, and the result will have fewer rows than ymat. However, if a matched set has several controls but no treated subject, then these controls will contribute to the estimate of the scale parameter, typically the median absolute pair difference.

## Note

The example reproduces Table 3 in Rosenbaum (2007).

## Author(s)

Paul R. Rosenbaum

## References

Rosenbaum, P. R. (2007) Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. Biometrics, 2007, 63, 456-464.

## Examples

```
data(tbmetaphase)
mscorev(tbmetaphase,trim=1)
```

mtm *DNA damage from exposure to chromium*

## Description

The data are from a study by Meibian et al. (2008) concerning possible damage to human DNA from occupational exposure to chromium. There were three matched groups, the control group (cmtm), a low exposure group (e2mtm) and a high exposure group (e1mtm). The exposed individuals all worked at a tannery where chromium was used in the tanning of leather. The highly exposed group (e1) worked at tanning leather. The low exposure group (e2) worked at the same tannery but did not tan leather. The reported values are the mean tail moment (mtm) of the comet assay, a measure of damage to DNA. High values of mtm indicate greater damage.

## Usage

```
data(mtm)
```

## Format

Each row of mtm is a matched set. The columns refer to the treatment groups mentioned in the description.

## Details

These data were used as an example of approximate evidence factors in Rosenbaum (2011). Under the null hypothesis H0 of no treatment effect, there are two approximately independent tests of H0 subject to different biases of nonrandom selection, specifically the comparison of 30 controls and 60 matched tannery workers, and the comparison of 30 low and 30 high exposure tannery workers. The two comparisons may be subjected to sensitivity analyses, say using senmv, and the results of these two analyses may be combined, for instance using Fisher's method of combining indepedent P-values. See the documentation for trucatedP or truncatedPbg for an example.

## Source

Meibian et al. (2008). Used as an example in Rosenbaum (2011).

## References

Meibian, Z., Zhijian, C., Qing, C. et al. (2008) Investigating DNA damage in tannery workers occupationally exposed to tivalent chromium using the comet assay. Mutation Research 654, 45-51.

Rosenbaum, P. R. (2011) Some approximate evidence factors in observational studies. Journal of the American Statistical Association, 2011, 106, 285-295.

## Examples

```
data(mtm)
```

`multrnks` *Approximate scores for ranks.*

## Description

Of limited interest to most users, this function is sometimes called by senmv. The function calculates the large sample approximation to a rank score transformation in Lemma 1, expression (9) of Rosenbaum (2011).

## Usage

```
multrnks(rk, m1 = 2, m2 = 2, m = 2)
```

## Arguments

| | |
|---|---|
| `rk` | A vector of ranks that may include average ranks for ties. |
| `m1` | One of three rank score parameters in Rosenbaum (2011), specifically m1 = underline(m). |
| `m2` | One of three rank score parameters in Rosenbaum (2011), specifically m2 = overline(m). |
| `m` | One of three rank score parameters in Rosenbaum (2011), specifically m = m. |

## Value

Vector of length(rk) containing the scores for the ranks in rk.

## Author(s)

Paul R. Rosenbaum

## References

Rosenbaum, P. R. (2011) A new u-statistic with superior design sensitivity in matched observational studies. Biometrics 67, 1017-1027.

## Examples

```
multrnks(1:10)
multrnks(1:10,m1=12,m2=20,m=20)
```

`newurks` *Approximate scores for ranks of row ranges.*

## Description

Of limited interest to most users, this function is sometimes called by senmv. The function calculates the ranges for each row of smat, scores their ranks using the large sample approximation to a rank score transformation in Lemma 1, expression (9) of Rosenbaum (2011), as implemented in the function multrnks, and multiplies the rows by these rank scores.

## Usage

```
newurks(smat, m1 = 2, m2 = 2, m = 2)
```

## Arguments

| | |
|---|---|
| `smat` | A matrix. When called by senmv, smat is a matrix produced by mscorev. |
| `m1` | One of three rank score parameters in Rosenbaum (2011), specifically m1 = underline(m). |
| `m2` | One of three rank score parameters in Rosenbaum (2011), specifically m2 = overline(m). |
| `m` | One of three rank score parameters in Rosenbaum (2011), specifically m = m. |

## Value

The function calculates the ranges for each row of smat, scores the ranks of the ranges using the large sample approximation to a rank score transformation in Lemma 1, expression (9) of Rosenbaum (2011), as implemented in the function multrnks, and multiplies the rows of smat by these rank scores, returning the result.

## Author(s)

Paul R. Rosenbaum

## References

Rosenbaum, P. R. (2011) A new u-statistic with superior design sensitivity in matched observational studies. Biometrics 67, 1017-1027.

## Examples

```
data(lead150)
  head(newurks(mscorev(lead150)))
```

`senmv` *Sensitivity analysis in observational studies using Huber-Maritz M-statistics.*

## Description

Computes the large sample approximation to the upper bound on the one sided P-value testing the null hypothesis of no treatment effect in a matched observational study with one or more controls matched to each treated subject. Uses Huber's M-statistics as test statistics in the sense proposed by Maritz (1979). The one-sided alternative hypothesis is that treatment increases the level of response.

The senmv function has as its default the use of Huber's (1981) psi function, which is similar in certain respects to a trimmed mean. By default, the trimming occurs (i.e., Huber's psi function becomes horizonal) at trim=2.5 times the median (i.e., lambda = 1/2) of the absolute pair differences; see Maritz (1979) for the paired case and see Rosenbaum (2007, section 4) for matched sets with one or more controls. With trim = 2.5, Huber's psi function is somewhat analogous to a trimmed mean that trims five percent from each tail. If trim is changed to 1, then the psi function is horizonal at 1 times the median (i.e., lambda=1/2) of the absolute pair differences and is somewhat analogous to a midmean.

method="t" performs the permutational t-test, the permutation distribution of a type of mean; see, for instance, Pitman (1937).

The senmv function can be used in the calculation of approximate evidence factors, as discussed and illustrated in the documentation for the truncatedP and truncatedPbg functions.

## Usage

```
senmv(y, gamma = 1, method = NULL, inner = 0, trim = 2.5,

 lambda = 1/2, tau = 0, TonT=FALSE)
```

## Arguments

| | |
|---|---|
| `y` | If y is a vector, then y is the vector of treated-minus-control pair differences in outcomes in n=length(y) matched pairs. If y is an n by J matrix, then: (i) the rows are n matched sets, (ii) the first column is the treated response in a set, columns 2 to J contain the responses of controls in the same matched set. If a matched set has fewer than J-1 controls, then NA's fill the empty spaces in y. The test is one-sided, testing no treatment effect against the alternative hypothesis that treated responses are higher than control responses. Replace y by -y to test against the alternative that treated responses are lower than control responses; see the mtm example. |
| `gamma` | gamma is the sensitivity parameter, gamma=1 for a randomization test, gamma>1 for sensitivity bounds. Use of gamma<1 will generate an error. This parameter gamma is denoted by the upper case Greek letter gamma in the cited literature, for instance Rosenbaum (2007). |
| `method` | If method is NULL, then the method is determined by the parameters, namely inner, trim, lambda, and TonT. If method is not NULL, then these parameters are set according to the selected method and stated values of the parameters are ignored. The default values of the parameters are equivalent to method="h".<br>(i) method = "h" (Huber) is the default and sets inner=0, trim=2.5, lambda = 1/2, TonT=FALSE. Its psi function levels off at 2.5 times the median (lambda = |

1/2) of the absolute pair differences. Method h is often a good choice in small samples with few pairs or sets. Method h is often a good choice when the number of controls in each matched set is 6 or more. For long tailed distributions, it may be better to use the same parameter values except with a lower value of trim, perhaps trim = 2 or trim = 1.

(ii) method = "i" (inner) often has better design sensitivity than method = "h". It has inner = 1/2, ignoring absolute pair differences smaller than half the median. It has inner=1/2, trim=2.5, lambda = 1/2, TonT=FALSE. Method i is often a good choice for a moderate to large sample of matched pairs, say 100 or more pairs. Method i is often a good choice for a moderate to large number of matched triples, say 100 or more. See Rosenbaum (2013) for comparisons of methods h and i.

(iii) method = "t" (permutational t-test) permutes the observations themselves, not scores derived from the observations. It has inner=0, trim=Inf, lambda = 1/2, TonT=TRUE. This is the only method that sets TonT=TRUE. The test statistic gives the mean over sets of the mean treated-minus-control difference within sets, that is, the usual estimate of the effect of the treatment on the treated. See the example for mercury using method="t" and the discussion of the TonT parameter.

`inner` Inner trimming to increase design sensitivity. See the discussion of lambda. Use of inner<0 or inner>trim will generate an error.

`trim` Outer trimming for resistance to outliers. Setting trim = Inf does no trimming. See the discussion of lambda.

`lambda` Observations are scaled by the lambda quantile of the absolute pair differences, defaulting to the median of all paired absolute differences; see Rosenbaum (2007, section 4.2) for a precise definition in the case of multiple controls. If the lambda quantile of the absolute pair differences is 0, then scaling by 0 is impossible and an error may result; the solution is to increase lambda. If qu is the lambda quantile, absolute pair differences smaller than inner*qu receive weight 0, absolute pair differences larger than qu*trim receive weight 1, and between inner*qu and trim*qu weights increase linearly from 0 to 1. Use of lambda<=0 or lambda>=1 will generate an error. If inner=0 and trim=Inf, then this results in the permutational t-test in which the observations themselves are permuted, and in this case lambda is not used. Taking lambda = .95 and trim = 1 is similar to trimming 5 percent of the pair differences.

`tau` If tau=0, senmv tests the null hypothesis of no treatment effect. If tau is not 0, senmv tests the null hypothesis that the treatment effect is an additive shift of tau against the alternative that the effect is larger than tau.

`TonT` TonT refers to the effect of the treatment on the treated. The default is TonT=FALSE. For all methods except method="t", TonT=FALSE. TonT=TRUE is mostly relevant to the permutational t-test, method="t", but it remains an option whenever is.null(method)=TRUE. If TonT=FALSE, then the total score in set (row) i is divided by the number ni of individuals in row i, as in expression (8) in Rosenbaum (2007). This division by ni has few consequences when every matched set has the same number of individuals, but when set sizes vary, dividing by ni is intended to increase efficiency by weighting inversely as the variance; see the discussion in section 4.2 of Rosenbaum (2007). If TonT=TRUE, then the division is by ni-1, not by ni, and there is a further division by the total number of matched sets to make it a type of mean. If TonT=TRUE, then the statistic is the mean over matched sets of the treated-score minus the mean-of-the-control-

scores within matched sets, so it is weighted to estimate the effect of the treatment on the treated. See the mercury example with method="t".

## Value

| | |
|---|---|
| `pval` | Approximate upper bound on the one-sided P-value. |
| `deviate` | Deviate that is compared to the upper tail of the standard Normal distribution to obtain the P-value. |
| `statistic` | Value of the test statistic. |
| `expectation` | Maximum null expectation of the test statistic for the given value of gamma. |
| `variance` | Among null distributions that yield the maximum expectation, variance is the maximum possible variance for the given value of gamma. See Rosenbaum (2007, Section 4) and Gastwirth, Krieger and Rosenbaum (2000). |

## Note

Example erpcp reproduces parts of Table 1 in Rosenbaum (2007), example tbmetaphase reproduces parts of section 4.3 in Rosenbaum (2007). Examples lead150 and lead250 reproduce parts of Table 1 of Rosenbaum (2013).

Example mtm almost reproduces parts of section 6 of Rosenbaum (2011). The second P-value bound differs slightly from that reported in the paper. The reason is that the paper used one scale factor, the median absolute pair difference, for both tests based on all pairs, whereas senmv applied to the last two columns of mtm uses a scale factor derived just from these two columns. There is no obvious reason to prefer one approach over the other, and the approach in the paper would require a separate R function. Note that in this example, y is -mtm because we test that the controls are low rather than high.

See the help file for mscorev to reproduce Table 3 in Rosenbaum (2007).

The mercury example illustrates the TonT parameter in a trivial case, because every treated subject has two controls. Note that the test statistic is a kind of mean when TonT=TRUE, but is merely a sum when TonT=FALSE. The weighting used when TonT=FALSE means the test statisitic uses information efficiently under certain models, but the test statistic itself is not an estimate of anything. Note the effect of setting tau = 2 is to reduce the test statistic by 2 when method="t".

## Author(s)

Paul R. Rosenbaum

## References

Main references:

Rosenbaum, P. R. (2007) Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. Biometrics, 2007, 63, 456-464. This paper is the main reference for the senmv function when weights are not used, m1=m2=m=1.

Rosenbaum, P. R. (2013) Impact of multiple matched controls on design sensitivity in observational studies. Biometrics, 2013, 69, 118-127. Evaluates the performance of the methods in the paper above, and in particular provides a basis for selecting the parameter values for senmv. In particular, this paper compares methods h, i, and t.

---

Additional references:

Cox, D. R. and Reid, N. Theory of the Design of Experiments. New York: Chapman and Hall/CRC. Chapter 2 discusses randomization inference, in particular an unmatched version of the permutation distributiom of the treated minus control difference in mean responses, or the two-sample (unmatched) permutational t-test. Although there is a large old literature on tests that permute the observations, this recent discussion is written in a modern style.

Fisher, R. A. (1935) Design of Experiments. Edinburgh: Oliver and Boyd. Chapter 3 contains an early, conceptual discussion of the permutation distribution of the mean or the permutational t-test.

Huber, P. (1981) Robust Statistics. New York: Wiley, 1981. Huber first proposed the use of m-statistics in 1964 in a paper in the Annals.

Maritz, J. S. (1979) Exact robust confidence intervals for location. Biometrika 1979, 66, 163-166. Proposed exact permutation tests using m-statistics that Maritz inverts to obtain exact confidence limits. The subtle aspect is the scaling which must be invariant to treatment assignment under the null hypothesis, so it differs from the scaling used by Huber.

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. Provides a general large sample approximation when matching with multiple controls, as used in Rosenbaum (2007, Section 4).

Pitman, E. J. G. (1937) Significance tests which may be applied to samples from any populations. JRSS-supplement (later called series B), 4, 119-130. An early technical discussion of the permutation distribution of the sample mean, or the permutational t-test.

Rosenbaum, P. R. (2010) Design of Observational Studies. New York: Springer 2010. Section 2.9 contains an elementary textbook discussion of Maritz's permutation distribution for m-statistics.

Rosenbaum, P. R. (2011) Some approximate evidence factors in observational studies. Journal of the American Statistical Association, 2011, 106, 285-295. The method described in this paper may be implemented using senmv. To do this, one uses senmv several times, combining the resulting one-sided P-value bounds, perhaps using Fisher's method for combining P-values. In the example of this paper, y is n x 3 for three groups in matched triples, and the paper uses Fisher's method to combine the P-value bounds from senmv(y) and senmv(y[,2:3]) for an appropriately defined y. See the mtm example in the documentation for truncatedP and truncatedPbg.

Welch, B. L. (1937) On the z-test in randomized blocks. Biometrika 29, 21-52. As in Pitman (1937) above, discusses permutation inference in which the responses are permuted, essentially a permutational F-test. Expresses causal effects as comparisons of potential responses under alternative treatments.

## Examples

```
# This example reproduces parts of Table 1 in Rosenbaum (2007).
data(erpcp)
senmv(erpcp,gamma=3,trim=1)
senmv(erpcp,gamma=2,trim=1)
senmv(erpcp,gamma=2,trim=1,tau=0.34)
senmv(erpcp,gamma=2,trim=1,tau=0.18)
senmv(erpcp,gamma=2,trim=1,tau=0.185)

# Example reproduces parts of sect. 4.3 in Rosenbaum (2007)
data(tbmetaphase)
senmv(tbmetaphase,gamma=2,trim=1)
senmv(tbmetaphase,gamma=2,trim=1,tau=0.94)
senmv(tbmetaphase,gamma=2,trim=1,tau=0.945)
```

```
# Example reproduces part of Table 1 in Rosenbaum (2013)
data(lead150)
senmv(lead150,gamma=2,trim=2)
data(lead250)
senmv(lead250,gamma=2,trim=2)

# Example reproduces parts of of Rosenbaum (2011).  See documentation for truncatedP.
data(mtm)
senmv(-mtm,gamma=11.7,trim=1)
senmv(-mtm[,2:3],gamma=2.1,trim=1)

# Illustrates method = "i"
data(mercury)
senmv(mercury,gamma=15)
senmv(mercury,gamma=15,method="i")

# Illustrates TonT=TRUE as in method="t".  See note above.
data(mercury)
mean(mercury[,1])-(mean(mercury[,2])+mean(mercury[,3]))/2
senmv(mercury,gamma=15,trim=Inf,TonT=TRUE)$statistic
senmv(mercury,gamma=15,method="t")$statistic
senmv(mercury,gamma=15,method="t",tau=1)$statistic
senmv(mercury,gamma=15,method="t",tau=2)$statistic
senmv(mercury,gamma=15,trim=Inf,TonT=FALSE)$statistic
```

---

`separable1v` *Asymptotic separable calculations internal to other functions.*

---

## Description

This general purpose function is internal to other functions, such as senmv in the sensitivitymv package. The function performs the asymptotic separable calculations described in Gastwirth, Krieger and Rosenbaum (2000), as used in section 4 of Rosenbaum (2007). The example is equivalent to senmv(lead150,gamma=2,method="t").

## Usage

```
separable1v(ymat, gamma = 1)
```

## Arguments

| | |
|---|---|
| `ymat` | ymat is a matrix whose rows are matched sets and whose columns are matched individuals. The first column describes treated individuals. Other columns describe controls. If matched sets contain variable numbers of controls, NAs fill in empty spaces in ymat; see the documentation for senmv. In senmv, the matrix ymat is created by mscorev. Instead, if there were no NAs and ranks within rows were used in ymat, then separable1v would perform a sensitivity analysis for the stratified Wilcoxon two-sample test. Applied directly to data, it performs a sensitivity analysis for the permutational t-test. |
| `gamma` | gamma is the value of the sensitivity parameter; see the documentation for the senmv function in the sensitivitymv package. One should use a value of gamma >= 1. |

## Value

| | |
|---|---|
| pval | Approximate upper bound on the one-sided P-value. |
| deviate | Deviate that is compared to the upper tail of the standard Normal distribution to obtain the P-value. |
| statistic | Value of the test statistic. |
| expectation | Maximum null expectation of the test statistic for the given value of gamma. |
| variance | Among null distributions that yield the maximum expectation, variance is the maximum possible variance for the given value of gamma. See Rosenbaum (2007, Section 4) and Gastwirth, Krieger and Rosenbaum (2000). |

## 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.

Rosenbaum, P. R. (2007) Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. Biometrics, 2007, 63, 456-464.

## Examples

```
data(lead150)
separable1v(lead150,gamma=2)
```

---

# tbmetaphase *Genetic damage from drugs used to treat TB*

---

## Description

This is a matched comparison of the effects of two drug sequences, namely HRZ and H2R2Z2, for the treatment of tuberculosis. HRZ is a higher dose sequence than H2R2Z2. The outcome is a measure of genetic damage, namely the frequency of aberrant metaphases two months after treatment. Individuals were matched for the frequency of aberrant metaphases before treatment. 15 individuals treated with HRZ are matched to 1 or 2 controls treated with H2R2Z2. Each row is one matched set. If a set is a pair, the third element in a row is NA. The data are originally from Rao, Gupta and Thomas (1991) and were used as an example in Rosenbaum (2007, Table 3). Data are used to illustrate the senmv function in the sensitivitymv package.

## Usage

```
data(tbmetaphase)
```

## Format

A data frame with 15 observations on the following 3 variables.

HRZ Aberrant metaphases for individual treated with HRZ.

H2R2Z2.1 Aberrant metaphases for first matched individual treated with H2R2Z2.

H2R2Z2.2 Aberrant metaphases for second matched individual treated with H2R2Z2. For matched pairs, this is NA.

## References

Rao, V. V. N. G., Gupta, E. V. V and Thomas, I. M. Chromosomal aberrations in tuberculosis patients before and after treatment with short-term chemotherapy. Mutation Research 1991, 259, 13-19.

Rosenbaum, P. R. Sensitivity analysis for m-estimates, tests and confidence intervals in matched observational studies. Biometrics, 2007, 63, 456-464.

## Examples

```
data(tbmetaphase)
```

---

`truncatedP` *Trucated product of P-values.*

---

## Description

Zaykin et al. (2002) proposed combining L=length(p) independent P-values p by taking the product of the P-values that are no larger than a truncation point trunc, namely w = prod(p^(p<=trunc)). Computes the one P-value for the combination w.

## Usage

```
truncatedP(p, trunc = 0.2)
```

## Arguments

| | |
|---|---|
| `p` | a vector of P-values that are either independent or stochastically larger than the uniform distribution on the cube (see Rosenbaum (2011, Definition 1)) |
| `trunc` | trucation point. Computes the product of all P-values in p that are no larger than trunc. For trunc = 1, performs Fisher's method of combining independent P-values. |

## Details

Zaykin et al. (2002) proposed combining L=length(p) independent P-values p by taking the product of the P-values that are no larger than a truncation point trunc, namely w = prod(p^(p<=trunc)). For trunc = 1, this is Fisher's method for combining independent P-values. The method also works for certain kinds of fairly inconsequential dependence; see Rosenbaum (2011, section 2).

trucatedP computes the one P-value for the combination w using the formula in Zaykin et al. (2002). The equivalent function truncatedPbg computes the exact same P-value for w using a binomial mixture of gamma distributions, as discussed by Hsu et al. (2013, section 3.1).

The truncated product or Fisher's method (trunc = 1) may be used for sensitivity analyses with evidence factors; see Rosenbaum (2011) and the mtm example below.

The truncated product with trunc < 1 is useful in combining P-value upper bounds produced by sensitivity analyses, for instance those produced by senmv. These upper bounds eventually approach 1 for larger values of the sensitivity parameter, and trunc < 1 eliminates these, often increasing power. See Hsu et al. (2013) for comparisons.

## Value

Returns the one P-value for the truncated product. For trunc < 1, the distribution is not continuous, but rather attaches positive probability to a P-value of 1.

## Note

Example mtm illustrates approximate evidence factors and almost reproduces parts of section 6 of Rosenbaum (2011) using the mtm data from Meibian et al. (2008). The second P-value bound differs slightly from that reported in the paper. The reason is that the paper used one scale factor, the median absolute pair difference, for both tests based on all pairs, whereas senmv applied to the last two columns of mtm uses a scale factor derived just from these two columns. There is no obvious reason to prefer one approach over the other, and the approach in the paper would require a separate R function. Note that in this example, y is -mtm because we test that the controls are low rather than high.

## Author(s)

Paul R. Rosenbaum

## References

Hsu, J. Y., Small, D. S. and Rosenbaum, P. R. (2013) Effect modification and design sensitivity in observational studies. Journal of the American Statistical Association, 108, 135-148.

Meibian, Z., Zhijian, C., Qing, C. et al. (2008) Investigating DNA damage in tannery workers occupationally exposed to tivalent chromium using the comet assay. Mutation Research 654, 45-51.

Rosenbaum, P. R. (2010) Evidence factors in observational studies. Biometrika, 97, 333-345.

Rosenbaum, P. R. (2011) Some approximate evidence factors in observational studies. Journal of the American Statistical Association, 106, 285-295.

Zaykin, D. V., Zhivotovsky, L. A., Westfall, P. H. and Weir, B. S. (2002) Truncated product method of combining P-values. Genetic Epidemiology, 22, 170-185.

Zhang, K., Small, D. S., Lorch, S., Srinivas, S. and Rosenbaum, P. R. (2011) Using split samples and evidence factors in an observational study of neonatal outcomes. Journal of the American Statistical Association, 106, 511-524.

## Examples

```
# Evidence factor example: see note above.
data(mtm)
senmv(-mtm,gamma=11.7,trim=1)
senmv(-mtm[,2:3],gamma=2.1,trim=1)
senmv(-mtm,gamma=12,trim=1)
senmv(-mtm[,2:3],gamma=3,trim=1)
truncatedP(c(0.05167572,0.1527849),trunc=1)
truncatedP(c(0.05167572,0.1527849),trunc=.2)
```

`truncatedPbg` *Trucated product of P-values using the mixture formula.*

## Description

Zaykin et al. (2002) proposed combining L=length(p) independent P-values p by taking the product of the P-values that are no larger than a truncation point trunc, namely w = prod(p^(p<=trunc)). Computes the one P-value for the combination w.

## Usage

```
truncatedPbg(p, trunc = 0.2)
```

## Arguments

| | |
|---|---|
| `p` | a vector of P-values that are either independent or stochastically larger than the uniform distribution on the cube (see Rosenbaum (2011, Definition 1)) |
| `trunc` | trucation point. Computes the product of all P-values in p that are no larger than trunc. For trunc = 1, performs Fisher's method of combining independent P-values. |

## Details

Zaykin et al. (2002) proposed combining L=length(p) independent P-values p by taking the product of the P-values that are no larger than a truncation point trunc, namely w = prod(p^(p<=trunc)). For trunc = 1, this is Fisher's method for combining independent P-values. The method also works for certain kinds of fairly inconsequential dependence; see Rosenbaum (2011, section 2).

trucatedP computes the one P-value for the combination w using the formula in Zaykin et al. (2002). The equivalent function truncatedPbg computes the exact same P-value for w using a binomial mixture of gamma distributions, as discussed by Hsu et al. (2013, section 3.1).

The truncated product or Fisher's method (trunc = 1) may be used for sensitivity analyses with evidence factors; see Rosenbaum (2011) and the mtm example below.

The truncated product with trunc < 1 is useful in combining P-value upper bounds produced by sensitivity analyses, for instance those produced by senmv. These upper bounds eventually approach 1 for larger values of the sensitivity parameter, and trunc < 1 eliminates these, often increasing power. See Hsu et al. (2013) for comparisons.

## Value

Returns the one P-value for the truncated product. For trunc < 1, the distribution is not continuous, but rather attaches positive probability to a P-value of 1.

## Note

Example mtm illustrates approximate evidence factors and almost reproduces parts of section 6 of Rosenbaum (2011) using the mtm data from Meibian et al. (2008). The second P-value bound differs slightly from that reported in the paper. The reason is that the paper used one scale factor, the median absolute pair difference, for both tests based on all pairs, whereas senmv applied to the last two columns of mtm uses a scale factor derived just from these two columns. There is no obvious reason to prefer one approach over the other, and the approach in the paper would require

a separate R function. Note that in this example, y is -mtm because we test that the controls are low rather than high.

## Author(s)

Paul R. Rosenbaum

## References

Hsu, J. Y., Small, D. S. and Rosenbaum, P. R. (2013) Effect modification and design sensitivity in observational studies. Journal of the American Statistical Association, 108, 135-148.

Meibian, Z., Zhijian, C., Qing, C. et al. (2008) Investigating DNA damage in tannery workers occupationally exposed to tivalent chromium using the comet assay. Mutation Research 654, 45-51.

Rosenbaum, P. R. (2010) Evidence factors in observational studies. Biometrika, 97, 333-345.

Rosenbaum, P. R. (2011) Some approximate evidence factors in observational studies. Journal of the American Statistical Association, 106, 285-295.

Zaykin, D. V., Zhivotovsky, L. A., Westfall, P. H. and Weir, B. S. (2002) Truncated product method of combining P-values. Genetic Epidemiology, 22, 170-185.

Zhang, K., Small, D. S., Lorch, S., Srinivas, S. and Rosenbaum, P. R. (2011) Using split samples and evidence factors in an observational study of neonatal outcomes. Journal of the American Statistical Association, 106, 511-524.

## Examples

```
# Evidence factor example: see note above.
data(mtm)
senmv(-mtm,gamma=11.7,trim=1)
senmv(-mtm[,2:3],gamma=2.1,trim=1)
senmv(-mtm,gamma=12,trim=1)
senmv(-mtm[,2:3],gamma=3,trim=1)
truncatedPbg(c(0.05167572,0.1527849),trunc=1)
truncatedPbg(c(0.05167572,0.1527849),trunc=.2)
```

# Index