

# Package ‘tailTransform’

May 8, 2026

**Type** Package

**Title** Symmetric Transformation of Tails for Plotting Differences

**Version** 2.0.0

**Author** Paul R. Rosenbaum [aut, cre]

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

**Description** When plotting treated-minus-control differences, after-minus-before changes, or difference-in-differences, the `ttrans()` function symmetrically transforms the positive and negative tails to aid plotting. The package includes an observational study with three control groups and an unaffected outcome; see Rosenbaum (2022) <[doi:10.1080/00031305.2022.2063944](https://doi.org/10.1080/00031305.2022.2063944)>.

**License** GPL-2

**Encoding** UTF-8

**LazyData** true

**Imports** stats, graphics

**Suggests** sensitivitymw, sensitivitymult

**Depends** R (>= 3.5.0)

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2025-03-22 21:00:02 UTC

## Contents

tailTransform-package	2
aHDL	3
boxplotTT	6
boxplotTTBlockDesign	7
boxplotTTlist	10
ttrans	12

<b>Index</b>	<b>14</b>
--------------	-----------

---

tailTransform-package *Symmetric Transformation of Tails for Plotting Differences*

---

## Description

When plotting treated-minus-control differences, after-minus-before changes, or difference-in-differences, the `ttrans()` function symmetrically transforms the positive and negative tails to aid plotting. The package includes an observational study with three control groups and an unaffected outcome; see Rosenbaum (2022) <doi:10.1080/00031305.2022.2063944>.

## Details

The DESCRIPTION file:

```
Package:      tailTransform
Type:        Package
Title:       Symmetric Transformation of Tails for Plotting Differences
Version:     2.0.0
Authors@R:   person(given = c("Paul", "R."), family = "Rosenbaum", role = c("aut", "cre"), email = "rosenbaum@wharton.u
Author:      Paul R. Rosenbaum [aut, cre]
Maintainer:  Paul R. Rosenbaum <rosenbaum@wharton.upenn.edu>
Description: When plotting treated-minus-control differences, after-minus-before changes, or difference-in-differences, the
License:     GPL-2
Encoding:    UTF-8
LazyData:   true
Imports:     stats, graphics
Suggests:   sensitivitymw, sensitivitymult
Depends:     R (>= 3.5.0)
```

Index of help topics:

aHDL	Alcohol and HDL Cholesterol: An Observational Study with 3 Control Groups
boxplotTT	Parallel Boxplots After Tail Transformation
boxplotTTBlockDesign	Parallel Boxplots After Tail Transformation For a Block Design
boxplotTTlist	Parallel Boxplots After Tail Transformation For a List
tailTransform-package	Symmetric Transformation of Tails for Plotting Differences
ttrans	Symmetric Tail Transformation of Differences for Graphical Display

The package contains five items: (i) a function `ttrans()` that symmetrically shortens the tails of pair differences for graphical display, (ii) three plotting functions, `boxplotTT()`, `boxplotTTlist()` and `boxplotBlockDesign()`, that aid in interpreting and displaying the transformed data, and (iii) an

observational study aHDL with three control groups and an unaffected outcome that are intended to reveal unmeasured confounding if it is present.<doi:10.1080/00031305.2022.2063944>

### Author(s)

Paul R. Rosenbaum [aut, cre]

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

### References

Rosenbaum, P. R. (2023) <doi:10.1111/biom.13558> Sensitivity analyses informed by tests for bias in observational studies. *Biometrics*, 9(1), 475-487.

Rosenbaum, P. R. (2022) <doi:10.1080/00031305.2022.2063944> A new transformation of treated-control matched-pair differences for graphical display. *American Statistician*, 76, 346-352.

### Examples

```
data(aHDL)
attach(aHDL)
d<-hdl[grpL=="D"]-hdl[grpL=="N"] # pair differences
tcks<-c(-100,-60,-40,-20,0,20,40,60,200)
boxplotTT(d,p=-1,qu=.95,tcks=tcks)
detach(aHDL)
rm(tcks,aHDL)
```

---

aHDL

*Alcohol and HDL Cholesterol: An Observational Study with 3 Control Groups*

---

### Description

A small observational study of light daily alcohol consumption and HDL cholesterol – so-called good cholesterol – derived from NHANES 2013-2014 and 2015-2016. There are 406 matched sets of four individuals, making 1624 individuals in total. Sets were matched for age, female and education in five ordered categories.

### Usage

```
data("aHDL")
```

### Format

A data frame with 1624 observations on the following 11 variables.

nh NHANES 2013-2014 is 1314, and NHANES 2015-2016 is 1516

SEQN NHANES ID number

age Age in years

female 1=female, 0=male

education 1 is <9th grade, 3 is high school, 5 is a BA degree

z 1=light almost daily alcohol, 0=little or no alcohol last year.

grp Treated group and control groups. Daily=light almost daily alcohol, Never=fewer than 12 drinks during entire life, Rarely=more than 12 drinks in life, but fewer than 12 in the past year, and never had a period of daily binge drinking, PastBinge = a past history of binge drinking on most days, but currently drinks once a week or less. For details, see Rosenbaum (2023, Appendix).

grpL Short labels for plotting formed as the first letters of grp. D < N < R < B

hdl HDL cholesterol level mg/dL

mmercury Methylmercury level ug/L

mset Matched set indicator, 1, 2, ..., 406. The 1624 observations are in 406 matched sets, each of size 4.

## Details

There is a debate about whether light daily alcohol consumption – a single glass of red wine – shortens or lengthens life. LoConte et al. (2018) emphasize that alcohol is a carcinogen. Suh et al. (1992) claim reduced cardiovascular mortality brought about by an increase in high density lipoprotein (HDL) cholesterol, the so-called good cholesterol. There is on-going debate about whether there are cardiovascular benefits, and if they exist, whether they are large enough to offset an increased risk of cancer. This example looks at a small corner of the larger debate, namely the effect on HDL cholesterol.

The example contains several attempts to detect unmeasured confounding bias, if present. There is a secondary outcome thought to be unaffected by alcohol consumption, namely methylmercury levels in the blood, likely an indicator of the consumption of fish, not of alcohol; see Pedersen et al. (1994) and WHO (2021). There are also three control groups, all with little present alcohol consumption, but with different uses of alcohol in the past; see the definition of variable grp above.

The appendix to Rosenbaum (2023) describes the data and matching in detail. It is used as an example in Rosenbaum (2022).

The help file for `boxplotTT()` applies the tail transformation to this example, reproducing a plot from Rosenbaum (2022).

## Source

US National Health and Nutrition Examination Survey (NHANES), 2013-2014 and 2015-2016. <[www.cdc.gov/nchs/nhanes](http://www.cdc.gov/nchs/nhanes)>

## References

LoConte, N. K., Brewster, A. M., Kaur, J. S., Merrill, J. K., and Alberg, A. J. (2018). Alcohol and cancer: a statement of the American Society of Clinical Oncology. *Journal of Clinical Oncology* 36, 83-93. <[doi:10.1200/JCO.2017.76.1155](https://doi.org/10.1200/JCO.2017.76.1155)>

Pedersen, G. A., Mortensen, G. K. and Larsen, E. H. (1994) Beverages as a source of toxic trace element intake. *Food Additives and Contaminants*, 11, 351–363. <[doi:10.1080/02652039409374234](https://doi.org/10.1080/02652039409374234)>

- Rosenbaum, P. R. (1987). The role of a second control group in an observational study. *Statistical Science*, 2, 292-306. <doi:10.1214/ss/1177013232>
- Rosenbaum, P. R. (1989). The role of known effects in observational studies. *Biometrics*, 45, 557-569. <doi:10.2307/2531497>
- Rosenbaum, P. R. (1989). On permutation tests for hidden biases in observational studies. *The Annals of Statistics*, 17, 643-653. <doi:10.1214/aos/1176347131>
- 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(507), 1145-1158 <doi:10.1080/01621459.2013.879261>
- Rosenbaum, P. R. (2023) <doi:10.1111/biom.13558> Sensitivity analyses informed by tests for bias in observational studies. *Biometrics*. 9(1), 475-487.
- Rosenbaum, P. R. (2022) <doi:10.1080/00031305.2022.2063944> A new transformation of treated-control matched-pair differences for graphical display. *American Statistician*, 76, 346-352.
- Suh, I., Shaten, B. J., Cutler, J. A., and Kuller, L. H. (1992). Alcohol use and mortality from coronary heart disease: the role of high-density lipoprotein cholesterol. *Annals of Internal Medicine* 116, 881-887. <doi:10.7326/0003-4819-116-11-881>
- World Health Organization (2021). Mercury and Health, <<https://www.who.int/news-room/fact-sheets/detail/mercury-and-health>>, (Accessed 30 August 2021).

## Examples

```
data(aHDL)
table(aHDL$grp,aHDL$grpL) # Short labels for plotting
boxplot(aHDL$age~aHDL$grp,xlab="Group",ylab="Age")
boxplot(aHDL$education~aHDL$grp,xlab="Group",ylab="Education")
table(aHDL$female,aHDL$grpL)
table(aHDL$z,aHDL$grpL)

# The sets were also matched for is.na(aHDL$mmercury), for use
# in Rosenbaum (2023). About half of the matched sets
# have values for mmercury.
table(is.na(aHDL$mmercury),aHDL$grp)

# Sensitivity analysis in Rosenbaum (2022); see
# also Rosenbaum (2014)
y<-t(matrix(aHDL$hdl,4,406))
y<-y[,c(1,3,2,4)]
colnames(y)<-c("D","N","R","B")
sensitivitymw::senmw(y,gamma=6,method="f")$pval
sensitivitymult::amplify(6,11)

# See also the informedSen package for additional analysis
```

---

`boxplotTT`*Parallel Boxplots After Tail Transformation*

---

**Description**

For a matrix of pair differences, plots one or more boxplots for differences after applying the same tail transformation to all differences. See the help file for `ttrans()` for information about the transformation. For groups of unequal sizes, use `boxplotTTlist()`.

**Usage**

```
boxplotTT(y, p = -1, qu = 0.95, tcks = NULL, ylab = "", xlab = "", main = "")
```

**Arguments**

<code>y</code>	A vector, matrix, or dataframe of differences to be transformed and plotted. Each column becomes a different boxplot after transformation.
<code>p</code>	The power of the transformation. See the help file for the <code>ttrans()</code> function.
<code>qu</code>	One number strictly between 0 and 1, commonly 0.9 or 0.95. Define beta to be the qu quantile of <code>abs(as.vector(y))</code> . Then, values between <code>-beta</code> and <code>beta</code> are not transformed.
<code>tcks</code>	A vector of untransformed values to become tick marks for the y axis after transformation. Although the y axis was transformed, it is labeled with corresponding untransformed values given by <code>tcks</code> . If <code>is.null(tcks)</code> , then the y axis has no tick marks.
<code>ylab</code>	Label for the y-axis.
<code>xlab</code>	Label for the x-axis.
<code>main</code>	Title for the plot.

**Details**

See the help file for `ttrans()` for an explanation of the transformation.

**Value**

A boxplot.

**Author(s)**

Paul R. Rosenbaum

**References**

Rosenbaum, P. R. (2022) <doi:10.1080/00031305.2022.2063944> A new transformation of treated-control matched-pair differences for graphical display. *American Statistician*, 76, 346-352.

**See Also**[ttrans](#)**Examples**

```

data(aHDL)
y<-t(matrix(aHDL$hdl,4,406))
y<-y[,c(1,3,2,4)]
colnames(y)<-c("D","N","R","B")

# 6 pairwise comparisons of 4 groups
o<-matrix(NA,dim(y)[1],6)
colnames(o)<-1:6
k<-0
for (i in 1:3)
  for (j in (i+1):4){
    k<-k+1
    colnames(o)[k]<-paste(colnames(y)[i],colnames(y)[j],sep="-")
    o[,k]<-y[,i]-y[,j]
  }
rm(i,j,k)

# Plotting tick marks. Remember, the transformation compresses
# extremes, so unequally spaced tick marks are usually needed.
tcks<-c(-100,-60,-40,-20,0,20,40,60,200)

# tails transformed by the p=-1 power (i.e., reciprocal)
boxplotTT(o,p=-1,qu=.95,tcks=tcks,ylab="HDL",
          xlab="Pairwise Comparisons of 4 Groups",
          main="HDL Differences in 4 Alcohol Groups")

```

---

boxplotTTBlockDesign *Parallel Boxplots After Tail Transformation For a Block Design*

---

**Description**

For an IxJ block design, a boxplot of column differences after transformation. See the help file for `ttrans()` for information about the transformation.

**Usage**

```

boxplotTTBlockDesign(bd, grp, p = -1, qu = 0.95, tcks = NULL, internal = TRUE,
  ylab = "", xlab = "", main = "", cex.axis = .8, cex.lab = .8,
  cex.main = .8, symZero=TRUE)

```

**Arguments**

<code>bd</code>	An $I \times J$ matrix or data.frame of outcomes with no missing outcomes.
<code>grp</code>	A vector of length $J = \text{number of columns of } bd$ . If <code>grp</code> is $1:J$ , then the plot contains $\text{choose}(J,2)$ boxplots for the differences of $\text{choose}(J,2)$ pairs of columns of <code>bd</code> , and these are labeled 1-2, 1-3, 2-3, .... If <code>grp</code> contains $J$ distinct short names, the same boxplots appear but labeled name1-name2, name1-name3,.... One-letter names look best. If $J=4$ and <code>grp</code> is (1,2,2,2) and <code>internal=FALSE</code> , then one boxplot appears, comparing column 1 to each of the columns labeled 2, using column 1 three times in comparison with columns 2, 3 and 4, so an $I \times 4$ design produces $3I$ differences. If $J=4$ and <code>grp</code> is (1,2,2,2) and <code>internal=TRUE</code> , then two boxplots appear. The first boxplot is the same as for <code>internal=FALSE</code> . The second boxplot displays $2I \times \text{choose}(J-1,2)$ differences, or $2I \times 3$ differences if <code>grp</code> is (1,2,2,2), namely $bd[,2]-bd[,3]$ , $bd[,3]-bd[,2]$ , $bd[,2]-bd[,4]$ , $bd[,4]-bd[,2]$ , $bd[,3]-bd[,4]$ , $bd[,4]-bd[,3]$ , that is, all pair differences among distinct columns labeled 2. See details. An error will result if all $J$ entries in <code>grp</code> are the same.
<code>p</code>	The power of the transformation. See the help file for the <code>ttrans()</code> function.
<code>qu</code>	One number strictly between 0 and 1, commonly 0.9 or 0.95. Define beta to be the <code>qu</code> quantile of $\text{abs}(\text{as.vector}(y))$ . Then, values between $-\text{beta}$ and $\text{beta}$ are not transformed.
<code>tcks</code>	A vector of untransformed values to become tick marks for the y axis after transformation. Although the y axis was transformed, it is labeled with corresponding untransformed values given by <code>tcks</code> . If $\text{is.null}(tcks)$ , then the y axis has no tick marks.
<code>internal</code>	If the $J$ entries in <code>grp</code> are all different, say <code>grp=(1,2,3,4)</code> , then changing <code>internal</code> has no effect. If one or more entries in <code>grp</code> are duplicated, as in <code>grp=(1,2,3,3)</code> , and if <code>internal=TRUE</code> , then three boxplots compare 1-2, 1-3 and 2-3, and a fourth boxplot displays the $2I$ symmetrized differences among controls labeled 3, namely $c(bd[,3]-bd[,4], bd[,4]-bd[,3])$ . If <code>grp=(1,2,3,3)</code> and <code>internal=FALSE</code> , then the fourth boxplot does not appear. See details.
<code>ylab</code>	Label for the y-axis.
<code>xlab</code>	Label for the x-axis.
<code>main</code>	Title for the plot. If <code>main=""</code> , then no title appears, and the sample size appears above each boxplot. If you prefer no title and no sample sizes, set <code>main=""</code> . Note that the sample size refers to the number of differences; see details.
<code>cex.lab</code>	Sets the label size for the boxplot using the standard graphics parameter <code>cex.lab</code> .
<code>cex.main</code>	Sets the title size for the boxplot using the standard graphics parameter <code>cex.main</code> .
<code>cex.axis</code>	Sets the axis size for the boxplot using the standard graphics parameter <code>cex.axis</code> .
<code>symZero</code>	If <code>TRUE</code> , makes the y-axis symmetric about zero. If <code>FALSE</code> , the maximum and minimum values in <code>ylist</code> determine the y-axis.

**Details**

See the help file for `ttrans()` for an explanation of the transformation.

If the columns of the  $I \times J$  block design `bd` represent different groups, say `grp=(1,2,...,J)`, then the  $\text{choose}(J,2)$  pairwise differences among pairs of columns are plotted after tail transformation.

Suppose that two or more columns represent the same group, so  $J > \text{length}(\text{unique}(\text{grp}))$ , as in  $\text{grp}(1,2,3,3,4,4)$  with a block size of  $J=6$  but just 4 groups. If  $\text{internal}=\text{FALSE}$ , then there is a boxplot plot of  $I$  differences comparing groups 1 and 2,  $\text{bd}[1]-\text{bd}[2]$ . There is a boxplot of  $2I$  differences comparing groups 1 and 3, namely  $c(\text{bd}[1]-\text{bd}[3], \text{bd}[1]-\text{bd}[4])$ . There is a boxplot of  $4I$  differences comparing groups 3 and 4.

In the same situation, but with  $\text{internal}=\text{TRUE}$ , the two columns for group 3 are compared, and so are the two columns for group 4. Both columns 5 and 6 each represent  $I$  individuals from group 4. A boxplot shows the variation in group 4 within blocks by plotting the  $2I$  symmetrized differences,  $c(\text{bd}[5]-\text{bd}[6], \text{bd}[6]-\text{bd}[5])$ . That boxplot is symmetric about zero, and shows how people in the same block differ by chance when they are in the same group. That boxplot is a benchmark for comparing two different groups.

For discussion and examples of symmetrized boxplots, see Ye et al. (2022, Figure 1) and Rosenbaum (2022, Figure 5).

### Value

Parallel boxplots after tail transformation.

### Author(s)

Paul R. Rosenbaum

### References

Rosenbaum, P. R. (2022) <doi:10.1080/00031305.2022.2063944> A new transformation of treated-control matched-pair differences for graphical display. *American Statistician*, 76, 346-352.

Ye, T., Small, D. S. and Rosenbaum, P. R. (2022) <doi:10.1214/22-AOAS1611> Dimensions, power and factors in an observational study of behavioral problems after physical abuse of children. *Annals of Applied Statistics*, 16, 2732-2754.

### See Also

[ttrans](#)

### Examples

```
# Makes Figure 5(ii) in Rosenbaum (2022).

data(aHDL)
y<-t(matrix(aHDL$hdl,4,406))
y<-y[,c(1,3,2,4)]
grp<-c("D","N","R","B")

# Figure 3 in Rosenbaum (2022)
boxplotTTBlockDesign(y,grp=grp,tcks=c(-200,-50,-20,0,20,50,200),
  ylab="HDL Difference",xlab="Group Comparisons")

# Same figure, different transformation, p=-2
boxplotTTBlockDesign(y,p=-2,grp=grp,tcks=c(-200,-50,-20,0,20,50,200),
  ylab="HDL Difference",xlab="Group Comparisons")
```

```

# Figure 5 in Rosenbaum (2022). The three control groups have been merged.
# Note that the C-C boxplot is perfectly symmetric about zero, and less
# dispersed than the T-C boxplot.
grp<-c("D", "C", "C", "C")
boxplotTTBlockDesign(y, grp=grp, tcks=c(-200, -50, -20, 0, 20, 50, 200),
  ylab="HDL Difference", xlab="Group Comparisons")

# The same figure can be produced explicitly using boxplotTTlist.
TC<-c(y[,1]-y[,2], y[,1]-y[,3], y[,1]-y[,4])
CC<-c(y[,2]-y[,3], y[,2]-y[,4], y[,3]-y[,4])
CC<-c(CC, -CC)

ylist<-list(TC=TC, CC=CC)

boxplotTTlist(ylist, tcks=c(-200, -50, -20, 0, 20, 50, 200), ylab="HDL Difference")

# More variations: Keep group B separate
grp<-c("D", "C", "C", "B")
boxplotTTBlockDesign(y, grp=grp, tcks=c(-50, -20, 0, 20, 50),
  symZero=FALSE, main=" ", p=1/2)
boxplotTTBlockDesign(y, grp=grp, tcks=c(-200, -50, -20, 0, 20, 50, 200),
  main="", p=-1, internal=FALSE)

```

---

boxplotTTlist

*Parallel Boxplots After Tail Transformation For a List*


---

## Description

For a list of vectors of pair differences, plots boxplots for differences after applying the same tail transformation to all differences. See the help file for `ttrans()` for information about the transformation.

## Usage

```

boxplotTTlist(ylist, p=-1, qu=.95, tcks=NULL, ylab="", xlab="", main="",
  cex.lab=.8, cex.main=.8, cex.axis=.8, symZero=TRUE)

```

## Arguments

<code>ylist</code>	A list of vectors of pair differences. An error will result if <code>ylist</code> is not a list. The vectors need not have the same lengths, and this makes <code>boxplotTTlist()</code> more flexible than <code>boxplotTT()</code> .
<code>p</code>	The power of the transformation. See the help file for the <code>ttrans()</code> function.
<code>qu</code>	One number strictly between 0 and 1, commonly 0.9 or 0.95. Define $\beta$ to be the $qu$ quantile of $\text{abs}(\text{as.vector}(y))$ . Then, values between $-\beta$ and $\beta$ are not transformed.

tcks	A vector of untransformed values to become tick marks for the y axis after transformation. Although the y axis was transformed, it is labeled with corresponding untransformed values given by tcks. If <code>is.null(tcks)</code> , then the y axis has no tick marks.
ylab	Label for the y-axis.
xlab	Label for the x-axis.
main	Title for the plot. If <code>main=""</code> , then no title appears, and the sample size appears above each boxplot. If you prefer no title and no sample sizes, set <code>main=""</code> .
cex.lab	Sets the label size for the boxplot using the standard graphics parameter <code>cex.lab</code> .
cex.main	Sets the title size for the boxplot using the standard graphics parameter <code>cex.main</code> .
cex.axis	Sets the axis size for the boxplot using the standard graphics parameter <code>cex.axis</code> .
symZero	If TRUE, makes the y-axis symmetric about zero. If FALSE, the maximum and minimum values in <code>ylist</code> determine the y-axis.

### Details

See the help file for `ttrans()` for an explanation of the transformation.

### Value

Parallel boxplots after tail transformation.

### Author(s)

Paul R. Rosenbaum

### References

Rosenbaum, P. R. (2022) <doi:10.1080/00031305.2022.2063944> A new transformation of treated-control matched-pair differences for graphical display. *American Statistician*, 76, 346-352.

Ye, T., Small, D. S. and Rosenbaum, P. R. (2022) <doi:10.1214/22-AOAS1611> Dimensions, power and factors in an observational study of behavioral problems after physical abuse of children. *Annals of Applied Statistics*, 16, 2732-2754.

### See Also

[ttrans](#)

### Examples

```
# Makes Figure 5(ii) in Rosenbaum (2022).
# See also Figure 1 in Ye et al. (2022).
```

```
data(aHDL)
y<-t(matrix(aHDL$hdl,4,406))
y<-y[,c(1,3,2,4)]
colnames(y)<-c("D", "N", "R", "B")
```

```

TC<-c(y[,1]-y[,2],y[,1]-y[,3],y[,1]-y[,4])
CC<-c(y[,2]-y[,3],y[,2]-y[,4],y[,3]-y[,4])
CC<-c(CC,-CC)

ylist<-list(TC=TC,CC=CC)

boxplotTlist(ylist,tcks=c(-80,-40,-20,0,20,40,80),ylab="HDL Difference")

```

---

ttrans

*Symmetric Tail Transformation of Differences for Graphical Display*


---

## Description

Performs a differentiable, strictly increasing, odd transformation of treated-minus-control pair differences  $d$ , or after-minus-before changes, or difference-in-differences. The transformation allows one to see the undistorted center of a distribution that contains extreme outliers, while also seeing the outliers. The transformation  $t(d)$  is: (i) odd, meaning  $t(d) = -t(-d)$ , so positive and negative values of  $d$  are transformed symmetrically, (ii) for some number  $\beta > 0$ , the transformation leaves  $d$  untouched between  $-\beta$  and  $\beta$ , so  $t(d) = d$  for  $-\beta < d < \beta$ , (iii) the transformation has derivative 1 at  $-\beta$  and  $\beta$ , so it is smooth at the point where the nonlinear transformation begins to take effect.

## Usage

```
ttrans(d, p = -1, qu = NULL, beta = NULL)
```

## Arguments

d	A vector of differences to be transformed.
p	The power to be used in the transformation of the tails, with $p=0$ being the log, as in the Box-Cox-Tukey transformation.
qu	If qu is specified, it is a number strictly between 0 and 1, commonly 0.9 or 0.95. Then $\beta$ is set to be the qu quantile of $\text{abs}(d)$ . If $\text{qu}=.95$ , then 95 percent of the differences in $d$ are not transformed. You must specify either qu or $\beta$ , and you must not specify both qu and $\beta$ .
beta	The value $\beta$ mentioned in the description. You must specify either qu or $\beta$ , and you must not specify both qu and $\beta$ .

## Details

Recall that  $\beta > 0$ . Let  $y$  be one difference in  $d$ . If  $y=0$ , then  $t(y)=y=0$ . If  $0 < y \leq \beta$ , then  $t(y) = y$ . If  $y > \beta$ , then a Box-Cox-Tukey power transformation is applied to  $y$ , with the transformation relocated and scaled so that  $t(y)$  has derivative 1 at  $\beta$ . If  $y < 0$ , then  $t(y) = -t(|y|)$ . Properties of the transformation are discussed in Rosenbaum (2022).

Although  $t(y)$  is nonlinear, it is exactly linear with slope 1 between  $-\beta$  and  $\beta$ , and  $t(y)$  is smooth with slope 1 at  $-\beta$  and  $\beta$ . The nonlinear aspect of the transformation is barely visible near  $-\beta$  and  $\beta$ .

If  $d$  is symmetric about zero, then the transformed values are also symmetric about zero. If there is no effect on the differences, in the sense that they are symmetric about zero, then the transformed differences also exhibit no effect.

The transformation does not alter Wilcoxon's signed rank statistic, or other signed rank statistics. Specifically, the transformation does not alter the ranks of  $|d|$ , and it does not alter  $\text{sign}(d)$ .

The  $p=-1$  reciprocal transformation has an upper and lower asymptote, so it limits the range of the  $d$ 's, but it shows outliers clearly.

### Value

A vector of transformed values of  $d$ .

### Author(s)

Paul R. Rosenbaum

### References

Box, George E. P. and David R. Cox. (1964) An analysis of transformations. *Journal of the Royal Statistical Society: Series B* 26, 211-243. <doi:10.1111/j.2517-6161.1964.tb00553.x>

Rosenbaum, P. R. (2022) <doi:10.1080/00031305.2022.2063944> A new transformation of treated-control matched-pair differences for graphical display. *American Statistician*, 76, 346-352.

Tukey, J. W. (1949). One degree of freedom for non-additivity. *Biometrics*, 5, 232-242. <doi:10.2307/3001938>

Tukey, J. W. (1957). On the comparative anatomy of transformations. *Annals of Mathematical Statistics*, 28, 602-632. <doi:10.1214/aoms/1177706875>

### Examples

```
data(aHDL)
attach(aHDL)
d<-hdl[grpL=="D"]-hdl[grpL=="N"] # pair differences
detach(aHDL)
oldpar<-par()
par(mfrow=c(1,2))
boxplot(d) # untransformed
boxplot(ttrans(d,qu=.95,p=-1)) # reciprocal transformation of tails

par(mfrow=c(1,1))
# Label the transformed vertical axis with untransformed values
# Add -beta and beta on the right axis
tcks<-c(-100,-60,-40,-20,0,20,40,60,200)
boxplotTT(d,p=-1,qu=.95,tcks=tcks)

par<-oldpar
rm(aHDL,d,oldpar)
```

# Index

- \* **Box-Cox transformation**
    - ttrans, [12](#)
  - \* **Causal effect**
    - boxplotTT, [6](#)
    - boxplotTTBlockDesign, [7](#)
    - boxplotTTlist, [10](#)
  - \* **Causal inference**
    - aHDL, [3](#)
  - \* **Matched pair difference**
    - boxplotTT, [6](#)
    - boxplotTTBlockDesign, [7](#)
    - boxplotTTlist, [10](#)
  - \* **Multiple control groups**
    - aHDL, [3](#)
  - \* **Negative control**
    - aHDL, [3](#)
  - \* **Observational study**
    - aHDL, [3](#)
  - \* **Placebo outcome**
    - aHDL, [3](#)
  - \* **Transformation**
    - boxplotTT, [6](#)
    - boxplotTTBlockDesign, [7](#)
    - boxplotTTlist, [10](#)
  - \* **Unaffected outcome**
    - aHDL, [3](#)
  - \* **datasets**
    - aHDL, [3](#)
  - \* **difference-in-differences**
    - ttrans, [12](#)
  - \* **hplot**
    - boxplotTT, [6](#)
    - boxplotTTBlockDesign, [7](#)
    - boxplotTTlist, [10](#)
  - \* **package**
    - tailTransform-package, [2](#)
  - \* **pair differences**
    - ttrans, [12](#)
  - \* **transformation**
    - ttrans, [12](#)
- aHDL, [3](#)
  - boxplotTT, [6](#)
  - boxplotTTBlockDesign, [7](#)
  - boxplotTTlist, [10](#)
  - tailTransform(tailTransform-package), [2](#)
  - tailTransform-package, [2](#)
  - ttrans, [7, 9, 11, 12](#)