Package 'SEAsic'

February 19, 2015

Title Score Equity Assessment- summary index computation

Version 0.1

Author Anne Corinne Huggins-Manley [aut, cre], Douglas Whitaker [aut]

Maintainer Anne Corinne Huggins-Manley <`ahuggins@coe.ufl.edu`>

Description This package conducts Score Equity Assessment (SEA; Dorans, 2004) by calculating and plotting multiple SEA indices as introduced by a variety of authors and summarized by Huggins and Penfield (2012).

Depends R (>= 3.1.1)

License GPL-3

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2014-11-19 16:46:34

R topics documented:

`adx` *Absolute Difference for Pairs - AD(x)*

Description

The absolute difference index, $AD(x)$, is the absolute value of the simple arithmetic difference between one subpopulation's conditional equated score, $y_j(x)$,and another subpopulation's conditional equated score, $y_{j'}(x)$. Formally,

$$AD(x) = \frac{\mid y_j(x) - y_{j'}(x) \mid}{\sigma_x},$$

where x is a score on the original (i.e., not equated) scale, and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered a pairwise, conditional invariance method. It was adapted by Huggins and Penfield (2012) from the analogous unconditional index, MAD, presented by Kolen and Brennan (2004). It provides practitioners with the magnitude of equated score differences between a pair of subpopulations at each level of the original scale.

Usage

```
adx(x, g1, g2, d, s, ymax, xlab, color)
```

Arguments

`x`	a column vector of scores on which the AD(x) is conditioned
`g1`	a column vector of equated scores based on a single subpopulation (aligned with elements in x)
`g2`	a column vector of equated scores based on a different single subpopulation (aligned with elements in x)
`d`	a scalar of the difference that matters
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized adx)
`ymax`	A maximum value for the y axis of the plot (default = 3 + the maximum AD(x) value)
`xlab`	A label for the x axis of the plot (default = Score Scale)
`color`	color of AD(x) line and points on plot (default = red)

Value

A data frame of AD(x) indices, conditioned on the score scale

A plot of the AD(x) indices in reference to the difference that matters

Author(s)

Anne Corinne Huggins-Manley

References

- Huggins, A.C., & Penfield, R.D. (2012). An NCME instructional module on population invariance in linking and equating. Educational Measurement: Issues and Practices, 31, 27-40.
- Kolen, M.J., & Brennan, R.L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). NY: Springer.

See Also

`madp`

Examples

```
#Unstandardized AD(x) for subpopulation 1 and subpopulation 2 in the example data set, ex.data
adx(x=ex.data[,1],g1=ex.data[,3],g2=ex.data[,4],d=.5)

#Unstandardized AD(x) for subpopulation 1 and subpopulation 2 in the example data set, ex.data,
#with adjustments to the maximum y-axis on the plot, a new xlabel, and points/line in blue.
adx(x=ex.data[,1],g1=ex.data[,3],g2=ex.data[,4],d=.5,ymax=2,xlab="AD(x)",color="blue")

#Unstandardized AD(x) for subpopulation 4 and subpopulation 5 in the example data set, ex.data
adx(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],d=.5)

#Standardized AD(x) for subpopulation 4 and subpopulation 5 in the example data set, ex.data
adx(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],d=.5,s=4.2)
```

`dx` *Difference for Pairs - D(x)*

Description

The difference index, $D(x)$, for pairs is the simple arithmetic difference between one subpopulation's conditional equated score, $y_j(x)$,and another subpopulation's conditional equated score, $y_{j'}(x)$. Formally,

$$D(x) = \frac{y_j(x) - y_{j'}(x)}{\sigma_x},$$

where x is a score on the original (i.e., not equated) scale and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered a pairwise, conditional invariance method. It was adapted by Huggins and Penfield (2012) from the analogous unconditional index, MD, presented by Kolen and Brennan (2004). It provides practitioners with the direction and magnitude of equated score differences between a pair of subpopulations at each level of the original scale.

Usage

```
dx(x, g1, g2, d, s, ymax, xlab, color)
```

Arguments

`x`	a column vector of scores on which the D(x) is conditioned
`g1`	a column vector of equated scores based on a single subpopulation (aligned with elements in x)
`g2`	a column vector of equated scores based on a different single subpopulation (aligned with elements in x)
`d`	a scalar of the positive difference that matters (negative of d will be applied as the negative difference that matters)
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized D(x))
`ymax`	A maximum value for the y axis of the plot (default = 3 + the maximum D(x) value)(negative of ymax will be applied to ymin of plot)
`xlab`	A label for the x axis of the plot (default = Score Scale)
`color`	color of D(x) line and points on plot (default = red)

Value

A data frame of D(x) indices, conditioned on the score scale

A plot of the D(x) indices in reference to the difference that matters

Author(s)

Anne Corinne Huggins-Manley

References

- Huggins, A.C., & Penfield, R.D. (2012). An NCME instructional module on population invariance in linking and equating. Educational Measurement: Issues and Practices, 31, 27-40.
- Kolen, M.J., & Brennan, R.L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). NY: Springer.

See Also

`mdp`

Examples

```
#Unstandardized D(x) for subpopulation 1 and subpopulation 2 in the example data set, ex.data
dx(x=ex.data[,1],g1=ex.data[,3],g2=ex.data[,4],d=.5)

#Unstandardized D(x) for subpopulation 1 and subpopulation 2 in the example data set, ex.data,
#with adjustments to the maximum y-axis on the plot, a new xlabel, and points/line in blue.
dx(x=ex.data[,1],g1=ex.data[,3],g2=ex.data[,4],d=.5,ymax=2,xlab="AD(x)",color="blue")

#Unstandardized D(x) for subpopulation 4 and subpopulation 5 in the example data set, ex.data
dx(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],d=.5)
```

```
#Standardized D(x) for subpopulation 4 and subpopulation 5 in the example data set, ex.data
dx(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],d=.5,s=4.2)
```

`emaxd` *Expected Maximum Difference*

Description

The expected maximum difference index, $EMAXD$, locates the maximum absolute difference between subpopulation equated scores, $y_j(x)$, and the equated scores based on the overall population, $y(x)$, at each score level x, and then takes the expectation of those maximum scores across score levels. Formally,

$$EMAXD = \frac{\sum_x P_x\{\max[\mid y_j(x) - y(x) \mid]\}}{\sigma_x},$$

where P_x represents a proportion of examinees based on the population distribution specified in argument f, and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered an omnibus, unconditional index that was originally presented by Dorans and Holland (2000). It provides practitioners with a summary of the maximum differences found between subpopulation and overall equated scores.

Usage

```
emaxd(x, o, g, n, f, s)
```

Arguments

`x`	a column vector of scores on which the rsd is conditioned
`o`	a column vector of equated scores based on the overall population (aligned with elements in x)
`g`	column vectors of equated scores based on various subpopulations (aligned with elements in x)
`n`	a scalar indicating the number of groups
`f`	a column vector of relative frequency associated with each raw score (can be based on either overall population or a subpopulation) (aligned with elements in x)
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized emaxd)

Value

expected maximum difference

Author(s)

Anne Corinne Huggins-Manley

References

- Dorans, N.J., & Holland, P.W. (2000). Population invariance and the equitability of tests: Theory and the linear case. Journal of Educational Measurement, 37, 281-306.

See Also

`maxd`

Examples

```
#Unstandardized EMAXD across subpopulation 1 and subpopulation 2 in the example data set, ex.data
emaxd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4]),
n=2,f=ex.data[,8])

#Unstandardized EMAXD across subpopulations 1 thru 5 in the example data set, ex.data
emaxd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
n=5,f=ex.data[,8])

#Standardized EMAXD across subpopulations 1 thru 5 in the example data set, ex.data
emaxd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
n=5,f=ex.data[,8],s=4.2)
```

`ex.data` *Example Data File*

Description

An example data set that has column 1 as x (i.e., unequated scale), column 2 as the equated scores based on the overall population, column 3 as the equated scores based on subpopulation 1, column 4 as the equated scores based on subpopulation 2, column 5 as the equated scores based on subpopulation 3, column 6 as the equated scores based on subpopulation 4, column 7 as the equated scores based on subpopulation 5, and column 8 as the frequency of examinees in the overall population scoring at each level of x.

Usage

```
ex.data
```

Format

```
'data.frame': 21 obs. of  8 variables:
 $ x          : num  0 1 2 3 4 5 6 7 8 9 ...
 $ overall    : num  0.5 1.51 2.56 3.59 4.61 5.62 6.65 7.69 8.7 9.71 ...
 $ g1         : num  0.47 1.72 3.1 3.99 5.2 5.98 7.03 8.01 8.97 9.61 ...
 $ g2         : num  0.61 1.52 2.46 3.54 4.71 5.6 6.64 7.7 8.81 9.72 ...
```

```
 $ g3           : num  0.53 1.61 2.89 3.62 4.81 5.23 6.64 7.21 8.75 9.81 ...
 $ g4           : num  0.12 0.98 1.46 1.99 2.47 3.42 4.58 5.12 5.99 6.78 ...
 $ g5           : num  1.54 2.72 3.46 4.21 5.94 ...
 $ relative_freq: num  0.001 0.001 0.001 0.01 0.02 0.04 0.052 0.065 0.09 0.12 ...
```

Value

An example data file (class: data.frame) of six sets of equated scores conditioned on unequated (i.e., x) scores

Author(s)

Anne Corinne Huggins-Manley

`madp` *Mean Absolute Difference for Pairs*

Description

The mean absolute difference index, MAD, is a summary of the conditional $AD(x)$ index, specifically the mean of absolute differences at each score level x. Formally,

$$MAD(x) = \frac{\sum_x P_x \mid y_j(x) - y_{j'}(x) \mid}{\sigma_x},$$

where $y_j(x)$ is an equated score based on subpopulation j, $y_{j'}(x)$ is an equated score based on subpopulation j', P_x represents a proportion of examinees based on the population distribution specified in argument f, and σ_x is the standard deviation of x scores for the (sub)population of interest. It is considered a pairwise, unconditional invariance method. It was originally presented by Kolen and Brennan (2004). It provides practitioners with a summary of the magnitude of equated score differences between two subpopulations.

Usage

```
madp(x, g1, g2, f, s)
```

Arguments

`x`	a column vector of scores on which the rsd is conditioned
`g1`	a column vector of equated scores based on a single subpopulation (aligned with elements in x)
`g2`	a column vector of equated scores based on a different single subpopulation (aligned with elements in x)
`f`	a column vector of relative frequency associated with each raw score (can be based on either overall population or a subpopulation) (aligned with elements in x)
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized madp)

Value

mean absolute difference

Note

The equally weighted version of this index (Kolen & Brennan, 2004) can be obtained by inputting an f vector consisting of identical elements that sum to 1. For example, using f=c(rep(.047619,21)) with the example data set, `ex.data`.

Author(s)

Anne Corinne Huggins-Manley

References

- Kolen, M.J., & Brennan, R.L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). NY: Springer.

See Also

`adx`

Examples

```
#Unstandardized MAD for subpopulation 1 and subpopulation 2 in the example data set, ex.data
madp(x=ex.data[,1],g1=ex.data[,3],g2=ex.data[,4],f=ex.data[,8])

#Unstandardized MAD for subpopulation 4 and subpopulation 5 in the example data set, ex.data
madp(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],f=ex.data[,8])

#Standardized MAD for subpopulation 4 and subpopulation 5 in the example data set, ex.data
madp(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],f=ex.data[,8],s=4.2)
```

`maxd` *Maximum Difference*

Description

The maximum difference index, $MAXD$, locates the maximum absolute difference between conditional subpopulation equated scores, $y_j(x)$, and the equated scores based on the overall population, $y(x)$. Formally,

$$MAXD = \frac{\max[\mid y_j(x) - y(x) \mid]}{\sigma_x},$$

where x is a score on the original (i.e., not equated) scale and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered an omnibus, conditional index. It was originally presented by Dorans and Holland (2000). It provides practitioners with the maximum difference found between any one subpopulation's equated scores and the equated scores based on the overall population, at each level of the original scale.

Usage

```
maxd(x, o, g, n, d, s, ymax, xlab, color)
```

Arguments

`x`	a column vector of scores on which the maximum difference is conditioned
`o`	a column vector of equated scores based on the overall population (aligned with elements in x)
`g`	column vectors of equated scores based on various subpopulations (aligned with elements in x)
`n`	a scalar indicating the number of groups
`d`	a scalar of the difference that matters
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized maxd)
`ymax`	A maximum value for the y axis of the plot (default = 3 + the maximum maxD value)
`xlab`	A label for the x axis of the plot (default = Score Scale)
`color`	color of maxD line and points on plot (default = red)

Value

A data frame of maximum difference indices, conditioned on the score scale

A plot of the maximum difference indices in reference to the difference that matters

Author(s)

Anne Corinne Huggins-Manley

References

- Dorans, N.J., & Holland, P.W. (2000). Population invariance and the equitability of tests: Theory and the linear case. Journal of Educational Measurement, 37, 281-306.

See Also

`emaxd`

Examples

```
#Unstandardized MAXD across subpopulation 1 and subpopulation 2 in the example data set, ex.data
maxd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4]),
n=2,d=.5)

#Unstandardized MAXD across subpopulations 1 thru 5 in the example data set, ex.data
maxd(x=ex.data[,1],o=ex.data[,2],
```

```
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
n=5,d=.5)

#Standardized MAXD across subpopulations 1 thru 5 in the example data set, ex.data
maxd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
n=5,d=.5,s=4.2)
```

`mdp` *Mean Difference for Pairs*

Description

The mean difference index (MD) for pairs is the unconditional version of $D(x)$. It is the expectation across x of the arithmetic differences between any $y_j(x)$ and $y_{j'}(x)$ equated scores. Formally,

$$MD = \frac{\sum_x P_x[y_j(x) - y_{j'}(x)]}{\sigma_x},$$

where P_x is the proportion of examinees scoring at x and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered a pairwise, unconditional index. It was originally presented by Kolen and Brennan (2004). It provides practitioners with a summary of the magnitude and direction of mean differences between equated scores based on two subpopulations.

Usage

```
mdp(x, g1, g2, f, s)
```

Arguments

`x`	a column vector of scores on which the rsd is conditioned
`g1`	a column vector of equated scores based on a single subpopulation (aligned with elements in x)
`g2`	a column vector of equated scores based on a different single subpopulation (aligned with elements in x)
`f`	a column vector of relative frequency associated with each raw score (can be based on either overall population or a subpopulation) (aligned with elements in x)
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized mean difference)

Value

mean difference

Note

The equally weighted version of this index (Kolen & Brennan, 2004) can be obtained by inputting an f vector consisting of identical elements that sum to 1. For example, using f=c(rep(.047619,21)) with the example data set, ex.data.

Author(s)

Anne Corinne Huggins-Manley

References

- Kolen, M.J., & Brennan, R.L. (2004). Test equating, scaling, and linking: Methods and practices (2nd ed.). NY: Springer.

See Also

`dx`

Examples

```
#Unstandardized MD for subpopulations 1 and 2 in the example data set, ex.data
mdp(x=ex.data[,1],g1=ex.data[,3],g2=ex.data[,4],f= ex.data[,8])

#Unstandardized MD for subpopulations 4 and 5 in the example data set, ex.data
mdp(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],f= ex.data[,8])

#Standardized MD for subpopulations 4 and 5 in the example data set, ex.data
mdp(x=ex.data[,1],g1=ex.data[,6],g2=ex.data[,7],f= ex.data[,8],s=4.2)
```

`pc` *Plot a Single or Multiple Conditional Indices*

Description

Function can be used to place a single or multiple conditional indices onto one plot, with the ability to alter the y-axis label, x-axis label, colors for each index, shapes of data points for each index, the placement of the legend and the limits of the y-axis.

Usage

```
pc(con, n, s, d, connames, ylab, xlab, colc, pchc, yleg, xleg, ylim)
```

Arguments

`con`	a concatenated series of data frames produced by SEAsic functions for (up to 6) conditional indices
`n`	scalar indicating the number of conditional indices in the plot
`s`	scalar representing the standard deviation of equated scores in the overall population, or on any (sub)population of interest (e.g., synthetic population) (default is 1, which should be used when plotting unstandardized indices)
`d`	unstandardized difference that matters
`connames`	a row vector of quoted names for the conditional indices
`ylab`	a quoted label for the y axis (default is "Equating Dependence")
`xlab`	a quoted label for the x axis (default is "Score Scale")
`colc`	a color designation for the conditional indices (default is a series of n rainbow colors)
`pchc`	a row vector of shape designation for the conditional index data points (default is filled circle)
`yleg`	a scalar indicating the value of the y axis at which the legend should sit (default is the maximum conditional index value for the first data frame in con)
`xleg`	a scalar indicating the value of the x axis at which the legend should sit (default is the minimum value of the score scale in the first data frame in con)
`ylim`	a row vector length of 2 with the minimum and maximum values for the y axis (default set at finite values in the plot)

Value

plot of multiple conditional indices

Note

- All indices must be measured on the same scale (e.g., all unstandardized indices).
- Plotting a single conditional index can also be done via functions to obtain the values of any conditional index, but there are more options to customize the plots via the pc function.

Author(s)

Anne Corinne Huggins-Manley

See Also

`pcu`

Examples

```
#Obtaining and plotting the unstandardized RSD(x) for all five subpopulations
#in the example data set, ex.data
rsd_g1 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],d=.5)
rsd_g2 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,4],d=.5)
rsd_g3 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,5],d=.5)
rsd_g4 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,6],d=.5)
rsd_g5 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5)

pc(con=c(rsd_g1,rsd_g2,rsd_g3,rsd_g4,rsd_g5),n=5,d=.5,
connames=c("RSD Group 1","RSD Group 2","RSD Group 3","RSD Group 4","RSD Group 5"),
ylim=c(0,4),yleg=4)

#Obtaining and plotting the standardized RSD(x) for all five subpopulations
#in the example data set, ex.data
srsd_g1 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],d=.5,s=4.2)
srsd_g2 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,4],d=.5,s=4.2)
srsd_g3 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,5],d=.5,s=4.2)
srsd_g4 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,6],d=.5,s=4.2)
srsd_g5 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5,s=4.2)

pc(con=c(srsd_g1,srsd_g2,srsd_g3,srsd_g4,srsd_g5),n=5,d=.5,
connames=c("RSD Group 1","RSD Group 2","RSD Group 3","RSD Group 4","RSD Group 5"),
s=4.2,ylim=c(0,2),yleg=2,ylab="Standardized Equating Dependence")

#Obtaining and plotting the unstandardized RMSD(x) for all five subpopulations
#in the example data set, ex.data
rmsd <- rmsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),w=c(.1,.2,.4,.2,.1),d=.5)

pc(con=c(rmsd),n=1,d=.5,connames=c("RMSD"),
ylim=c(0,4),yleg=4,ylab="Root Mean Square Differences")
```

---

pcu *Plot Multiple Conditional/Unconditional Indices*

Description

Function can be used to place multiple conditional indices and their unconditional counterparts onto one plot, with the ability to alter the y-axis label, x-axis label, colors for each pair of indices, shapes of data points for each pair of indices, the placement of the legend and the limits of the y-axis.

Usage

```
pcu(con, u, s, d, connames, unames, ylab, xlab, colc, colu, pchc, ltyu, yleg,
  xleg, ylim)
```

Arguments

`con`	a concatenated series of data frames produced by SEAsic functions for (up to 6) conditional indices
`u`	a concatenated series of scalars representing (up to 6) unconditional indices, which can be entered directly or called in from previous SEAsic estimation of such indices
`s`	scalar representing the standard deviation of equated scores in the overall population, or on any (sub)population of interest (e.g., synthetic population) (default is 1, which should be used when plotting unstandardized indices)
`d`	unstandardized difference that matters
`connames`	a row vector of quoted names for the conditional indices
`unames`	a row vecotr of quoted names for the unconditional indices
`ylab`	a quoted label for the y axis (default is "Equating Dependence")
`xlab`	a quoted label for the x axis (default is "Score Scale")
`colc`	a color designation for the conditional indices (default is a series of rainbow colors the length of u)
`colu`	a color designation for the unconditional index lines (default is a series of rainbow colors matching colc)
`pchc`	a row vector of shape designation for the conditional index data points (default is filled circle)
`ltyu`	a row vector of line types for the unconditional index lines (default is straight line)
`yleg`	a scalar indicating the value of the y axis at which the legend should sit (default is the maximum conditional index value for the first data frame in con)
`xleg`	a scalar indicating the value of the x axis at which the legend should sit (default is the minimum value of the score scale in the first data frame in con)
`ylim`	a row vector length of 2 with the minimum and maximum values for the y axis (default set at finite values in the plot)

Value

plot of multiple conditional and unconditional indices

Note

All indices must be measured on the same scale (e.g., all unstandardized indices), and all conditional indices must have an unconditional counterpart in the plot.

Author(s)

Anne Corinne Huggins-Manley

See Also

`pc`

Examples

```
#Obtaining and plotting the unstandardized RSD(x) and RESD for all five subpopulations
#in the example data set, ex.data
rsd_g1 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],d=.5)
rsd_g2 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,4],d=.5)
rsd_g3 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,5],d=.5)
rsd_g4 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,6],d=.5)
rsd_g5 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5)
resd_g1 <- resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],f=ex.data[,8])
resd_g2 <- resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,4],f=ex.data[,8])
resd_g3 <- resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,5],f=ex.data[,8])
resd_g4 <- resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,6],f=ex.data[,8])
resd_g5 <- resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],f=ex.data[,8])

pcu(con=c(rsd_g1,rsd_g2,rsd_g3,rsd_g4,rsd_g5),
u=c(resd_g1,resd_g2,resd_g3,resd_g4,resd_g5),d=.5,
connames=c("RSD Group 1","RSD Group 2","RSD Group 3","RSD Group 4","RSD Group 5"),
unames=c("RESD Group 1","RESD Group 2","RESD Group 3","RESD Group 4","RESD Group 5"),
ylim=c(0,8),yleg=8)

#Obtaining and plotting the standardized RSD(x) and RESD for all five subpopulations
#in the example data set, ex.data
srsd_g1 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],d=.5,s=4.2)
srsd_g2 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,4],d=.5,s=4.2)
srsd_g3 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,5],d=.5,s=4.2)
srsd_g4 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,6],d=.5,s=4.2)
srsd_g5 <- rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5,s=4.2)
sresd_g1<-resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],f=ex.data[,8],s=4.2)
sresd_g2<-resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,4],f=ex.data[,8],s=4.2)
sresd_g3<-resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,5],f=ex.data[,8],s=4.2)
sresd_g4<-resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,6],f=ex.data[,8],s=4.2)
sresd_g5<-resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],f=ex.data[,8],s=4.2)

pcu(con=c(srsd_g1,srsd_g2,srsd_g3,srsd_g4,srsd_g5),
u=c(sresd_g1,sresd_g2,sresd_g3,sresd_g4,sresd_g5),d=.5,s=4.2,
connames=c("RSD Group 1","RSD Group 2","RSD Group 3","RSD Group 4","RSD Group 5"),
unames=c("RESD Group 1","RESD Group 2","RESD Group 3","RESD Group 4","RESD Group 5"),
ylim=c(0,2),yleg=2,ylab="Standardized Equating Dependence")

#Obtaining and plotting the unstandardized RMSD(x) and REMSD for all five subpopulations
#in the example data set, ex.data
rmsd <- rmsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
w=c(.1,.2,.4,.2,.1),d=.5)
remsd <- remsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
f=ex.data[,8],w=c(.1,.2,.4,.2,.1))

pcu(con=c(rmsd),u=c(remsd),d=.5,connames=c("RMSD"),unames=c("REMSD"),ylim=c(0,4),yleg=4)
```

`remsd` *Root Expected Mean Square Difference*

Description

The root expected mean square difference index ($REMSD$) is a summary index of the weighted differences between each subpopulation equated score, $y_j(x)$, and the equated score based on the overall population, $y(x)$. Formally,

$$REMSD = \frac{\sqrt{\sum_x P_x\{\sum_j w_j[y_j(x) - y(x)]^2\}}}{\sigma_x,}$$

where w_j is a subpopulation weight, x is a score on the original (i.e., unequated) scale, P_x is the proportion of examinees scoring at x and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered an omnibus, unconditional index. It was originally presented by Dorans and Holland (2000). It provides practitioners with a summary of the magnitude of weighted differences between subpopulation equated scores and equated scores based on the overall population.

Usage

```
remsd(x, o, g, f, s, w)
```

Arguments

`x`	a column vector of scores on which the rmsd is conditioned
`o`	a column vector of equated scores based on the overall population (aligned with elements in x)
`g`	column vectors of equated scores based on various subpopulations (aligned with elements in x)
`f`	a column vector of relative frequency associated with each raw score (can be based on either overall population or a subpopulation) (aligned with elements in x)
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized remsd)
`w`	A row vector of weights for subpopulations 1 thru n (length = number of groups)

Value

root expected mean square difference

Note

The equally weighted version of this index (Kolen & Brennan, 2004) can be obtained by inputting a w vector consisting of identical elements that sum to 1. See example 1 below.

Author(s)

Anne Corinne Huggins-Manley

References

- Dorans, N.J., & Holland, P.W. (2000). Population invariance and the equitability of tests: Theory and the linear case. Journal of Educational Measurement, 37, 281-306.

See Also

`rmsd`

Examples

```
#Unstandardized REMSD for subpopulations 1 and 2 in the example data set, ex.data,
#assuming equal weights for the subpopulations
remsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4]),f=ex.data[,8],w=c(.5,.5))

#Unstandardized REMSD for all five subpopulations in the example data set, ex.data
remsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
f=ex.data[,8],w=c(.1,.2,.4,.2,.1))

#Standardized REMSD for all five subpopulations in the example data set, ex.data
remsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
f=ex.data[,8],w=c(.1,.2,.4,.2,.1),s=4.2)
```

`resd` *Root Expected Square Difference*

Description

The root expected square difference index ($RESD_j$) is a summary index of the weighted differences between a single subpopulation's equated scores, $y_j(x)$, and the equated scores based on the overall population, $y(x)$. Formally,

$$RESD_j = \frac{\sqrt{\sum_x P_x\{[y_j(x) - y(x)]^2\}}}{\sigma_x,}$$

where x is a score on the original (i.e., unequated) scale, P_x is the proportion of examinees scoring at x and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered a group-to-overall, unconditional index. It was originally presented by Yang (2004). It provides practitioners with a summary of the magnitude of differences between a single subpopulation's equated scores and equated scores based on the overall population.

Usage

```
resd(x, o, g, f, s)
```

Arguments

`x`	a column vector of scores on which the rsd is conditioned
`o`	a column vector of equated scores based on the overall population (aligned with elements in x)
`g`	a column vector of equated scores based on a single subpopulation (aligned with elements in x)
`f`	a column vector of relative frequency associated with each raw score (can be based on either overall population or a subpopulation) (aligned with elements in x)
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized resd)

Value

root expected square difference

Author(s)

Anne Corinne Huggins-Manley

References

- Yang, W.L. (2004). Sensitivity of linkings between AP multiple-choice scores and composite scores to geographical region: An illustration of checking for population invariance. Journal of Educational Measurement, 41, 33-41.

See Also

`rsd`

Examples

```
#Unstandardized RESD for subpopulation 1 in the example data set, ex.data
resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],f=ex.data[,8])

#Unstandardized RESD for subpopulation 5 in the example data set, ex.data
resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],f=ex.data[,8])

#Standardized RESD for subpopulation 5 in the example data set, ex.data
resd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],f=ex.data[,8],s=4.2)
```

`rmsd` *Root Mean Square Difference*

Description

The root mean square difference index ($RMSD$) is a conditional index of the weighted differences between each subpopulation equated score, $y_j(x)$, and the equated score based on the overall population, $y(x)$. Formally,

$$RMSD = \frac{\sqrt{\sum_j w_j [y_j(x) - y(x)]^2}}{\sigma_x},$$

where w_j is a subpopulation weight, x is a score on the original (i.e., unequated) scale, and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered an omnibus, conditional index. It was originally presented by Dorans and Holland (2000). It provides practitioners with the magnitude of weighted differences between subpopulation equated scores and equated scores based on the overall population.

Usage

```
rmsd(x, o, g, w, d, s, ymax, xlab, color)
```

Arguments

`x`	a column vector of scores on which the rmsd is conditioned
`o`	a column vector of equated scores based on the overall population (aligned with elements in x)
`g`	column vectors of equated scores based on various subpopulations (aligned with elements in x)
`w`	a row vector of weights for subpopulations 1 thru n (length = number of groups)
`d`	a scalar of the difference that matters
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized rmsd)
`ymax`	A maximum value for the y axis of the RMSD graph (default = 3 + the maximum RMSD value)
`xlab`	A label for the x axis of the RMSD graph (default = Score Scale)
`color`	color of RMSD line and points on plot (default = red)

Value

A data frame of RMSD indices, conditioned on the score scale

A plot of the RMSD in reference to the difference that matters

Note

The equally weighted version of this index (Kolen & Brennan, 2004) can be obtained by inputting a w vector consisting of identical elements that sum to 1. See example 1 below.

Author(s)

Anne Corinne Huggins-Manley

References

- Dorans, N.J., & Holland, P.W. (2000). Population invariance and the equitability of tests: Theory and the linear case. Journal of Educational Measurement, 37, 281-306.

See Also

`remsd`

Examples

```
#Unstandardized RMSD for subpopulations 1 and 2 in the example data set, ex.data,
#assuming equal weights for the subpopulations
rmsd(x=ex.data[,1],o=ex.data[,2],g=c(ex.data[,3],ex.data[,4]),w=c(.5,.5),d=.5)

#Unstandardized RMSD for all five subpopulations in the example data set, ex.data
rmsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
w=c(.1,.2,.4,.2,.1),d=.5)

#Unstandardized RMSD for all five subpopulations in the example data set, ex.data,
#with adjustments to the maximum y-axis on the plot, a new xlabel,
#and points/line in blue
rmsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
w=c(.1,.2,.4,.2,.1),d=.5,ymax=3,xlab="Old Form",color="blue")

#Standardized RMSD for all five subpopulations in the example data set, ex.data
rmsd(x=ex.data[,1],o=ex.data[,2],
g=c(ex.data[,3],ex.data[,4],ex.data[,5],ex.data[,6],ex.data[,7]),
w=c(.1,.2,.4,.2,.1),d=.5,s=4.2)
```

`rsd` *Root Square Difference*

Description

The root square difference index (RSD) is a conditional index of the absolute differences between a single subpopulation j's equated score, $y_j(x)$, and the equated score based on the overall population, $y(x)$. Formally,

$$RSD_j(x) = \frac{\mid y_j(x) - y(x) \mid}{\sigma_x},$$

where x is a score on the original (i.e., unequated) scale, and σ_x is the standard deviation of x scores in the (sub)population of interest. It is considered a group-to-overall, conditional index. It was adapted from Yang's $RESD_j$ (2004) by Huggins and Penfield (2012). It provides practitioners with the magnitude of differences between a single subpopulation's equated scores and the equated score based on the overall population, at each score level on the original score scale.

Usage

```
rsd(x, o, g, d, s, ymax, xlab, color)
```

Arguments

`x`	a column vector of scores on which the rsd is conditioned
`o`	a column vector of equated scores based on the overall population (aligned with elements in x)
`g`	a column vector of equated scores based on a single subpopulation (aligned with elements in x)
`d`	a scalar of the difference that matters
`s`	a scalar representing the standard deviation of x for any (sub)population of interest (e.g., synthetic population) (default is 1, which leads to calculation of the unstandardized rsd)
`ymax`	A maximum value for the y axis of the plot (default = 3 + the maximum RSD value)
`xlab`	A label for the x axis of the plot (default = Score Scale)
`color`	of RSD line and points on plot (default = red)

Value

A data frame of root square difference indices, conditioned on the score scale

A plot of the RSD in reference to the difference that matters

Author(s)

Anne Corinne Huggins-Manley

References

- Huggins, A.C., & Penfield, R.D. (2012). An NCME instructional module on population invariance in linking and equating. Educational Measurement: Issues and Practices, 31, 27-40.
- Yang, W.L. (2004). Sensitivity of linkings between AP multiple-choice scores and composite scores to geographical region: An illustration of checking for population invariance. Journal of Educational Measurement, 41, 33-41.

See Also

`resd`

Examples

```
#Unstandardized RSD for subpopulation 1 in the example data set, ex.data
rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,3],d=.5)

#Unstandardized RSD for subpopulation 5 in the example data set, ex.data
rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5)

#Unstandardized RSD for subpopulation 5 in the example data set, ex.data
#with adjustments to the maximum y-axis on the plot, a new x label,
#and points/line in green
rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5,ymax=3,xlab="Old Form",color="green")

#Standardized RSD for subpopulation 5 in the example data set, ex.data
rsd(x=ex.data[,1],o=ex.data[,2],g=ex.data[,7],d=.5,s=4.2)
```

Index