Package ‘wINEQ’

July 22, 2025

Title Inequality Measures for Weighted Data

Version 1.2.1

Description Computes inequality measures of a given variable taking into account weights. Suitable for ratio, interval and ordered scale. Includes Gini, Theil, Leti index, Palma ratio, 20:20 ratio, Allison and Foster index, Jenkins index, Cowell and Flechaire index, Abul Naga and Yalcin index, Apouey index, Blair and Lacy index. Bootstrap provides distribution of inequality measures enabling significance tests.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.2

Depends dplyr, sampling, stats, R (>= 2.10)

LazyData true

NeedsCompilation no

Author Sebastian Wójcik [aut, cre] (ORCID: <https://orcid.org/0000-0003-2425-9626>),
Agnieszka Giemza [aut],
Katarzyna Machowska [aut],
Jarosław Napora [aut]

Maintainer Sebastian Wójcik <S.Wojcik@stat.gov.pl>

Repository CRAN

Date/Publication 2024-09-09 11:20:08 UTC

Contents

AF 2
AN_Y 3
Apouey 4
Atkinson 5
BL 6
CoefVar 8
Entropy 9
Gini 10

Hoover 11
ineq.weighted 12
ineq.weighted.boot 14
Jenkins 16
Kolm 17
Leti 18
LowerSum 20
medianf 20
Palma 21
Prop20_20 22
Quantile 23
RicciSchutz 24
Theil_L 25
Theil_T 26
Tourism 28
Well_being 29

Index **31**

`AF` *Allison and Foster index*

Description

Computes Allison and Foster inequality measure of a given variable taking into account weights.

Usage

```
AF(X, W = rep(1, length(X)), norm = TRUE)
```

Arguments

`X`	is a data vector (numeric or ordered factor)
`W`	is a vector of weights
`norm`	(logical). If TRUE (default) then index is divided by a maximum possible value which is a difference between maximum and minimum of X

Details

Let $c = (c_1, ..., c_n)$ be the vector of categories in increasing order, m be the median category and p_i be a share of i-th category. The following index was proposed by Allison and Foster (2004):

$$AF = \frac{\sum_{i=m}^{n} c_i p_i}{\sum_{i=m}^{n} p_i} - \frac{\sum_{i=1}^{m-1} c_i p_i}{\sum_{i=1}^{m-1} p_i}$$

Note that above formula is valid only for numerical values. Thus, in order to compute AF for ordered factor, X is converted to numerical variable.

Value

The value of Allison and Foster coefficient.

References

Allison R. A., Foster J E.: (2004) Measuring health inequality using qualitative data, Journal of Health Economics

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
AF(X)
AF(X,W)

data(Well_being)
# Allison and Foster index for health assessment with sample weights
X=Well_being$V11
W=Well_being$Weight
AF(X,W)
```

AN_Y *Abul Naga and Yalcin index*

Description

Computes Abul Naga and Yalcin inequality measure of a given variable taking into account weights.

Usage

```
AN_Y(X, W = rep(1, length(X)), a = 1, b = 1)
```

Arguments

X	is a data vector (numeric or ordered factor)
W	is a vector of weights
a	is a positive parameter. See more in details
b	is a positive parameter. See more in details

Details

Let m be the median category, n be the number of categories and P_i be the cumulative distribution of i-th category. The following index with respect to the parameters a and b was proposed by Abul Naga and Yalcin (2008):

$$I = \frac{a\sum_{i<m}^{n} P_i - b\sum_{i\geq m}^{n} P_i + b(n+1-m)}{0.5(a(m-1)+b(n-m))}$$

Value

The value of Abul Naga and Yalcin coefficient.

References

Ramses H. Abul Naga and Tarik Yalcin: (2008) Inequality Measurement for ordered response health data, Journal of Health Economics 27(6);

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
AN_Y(X)
AN_Y(X,W)

data(Well_being)
# Abul Naga and Yalcin index for health assessment with sample weights
X=Well_being$V1
W=Well_being$Weight
AN_Y(X,W)
```

Apouey *Apouey index*

Description

Computes Apouey inequality measure of a given variable taking into account weights.

Usage

```
Apouey(
  X,
  W = rep(1, length(X)),
  a = 2/(1 - length(W[!is.na(W) & !is.na(X)])),
  b = length(W[!is.na(W) & !is.na(X)])/(length(W[!is.na(W) & !is.na(X)]) - 1)
)
```

Arguments

X	is a data vector (numeric or ordered factor)
W	is a vector of weights
a	is a positive parameter. See more in details
b	is a real parameter. See more in details

Details

Let m be the median category, n will be the number of categories and P_i be the cumulative distribution of i-th category. The following index was proposed by Apouey (2007):

$$I = \alpha(\sum_{i \geq m}^{n} P_i - \sum_{i<m}^{n} P_i + m - \frac{n}{2} - 1) + \beta$$

where α and β are given parameters with default values $\alpha = \frac{2}{1-n}$ and $\beta = \frac{n}{n-1}$.

Value

The value of Apouey coefficient.

References

Apouey B.: (2007) Measuring health polarization with self-assessed health data, Health Economics 16; 875-894.

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Apouey(X,a=2,b=2)
Apouey(X,W,a=2,b=2)

data(Well_being)
# Apouey index for health assessment with sample weights
X=Well_being$V1
W=Well_being$Weight
Apouey(X,W,a=2,b=2)
```

`Atkinson` *Atkinson index*

Description

Computes Atkinson inequality measure of a given variable taking into account weights.

Usage

```
Atkinson(X, W = rep(1, length(X)), e = 1)
```

Arguments

`X`	is a data vector
`W`	is a vector of weights
`e`	is a coefficient of aversion to inequality, by default 1

Details

Atkinson coefficient with respect to parameter ϵ is given by

$$1 - \frac{1}{\mu}(\frac{1}{n}\sum_{i=1}^{n} x_i^{1-\epsilon})^{\frac{1}{1-\epsilon}}$$

for $\epsilon \neq 1$ and

$$1 - \frac{1}{\mu}(\prod_{i=1}^{n} x_i)^{\frac{1}{n}}$$

for $\epsilon = 1$.

Value

The value of Atkinson coefficient.

References

Atkinson A. B.: (1970) On the measurement of inequality, Journal of Economic Theory

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Atkinson(X)
Atkinson(X,W)

data(Tourism)
# Atkinson index for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Atkinson(X,W)
```

BL *Blair and Lacy index*

Description

Computes Blair and Lacy inequality measure of a given variable taking into account weights.

Usage

```
BL(X, W = rep(1, length(X)), withsqrt = FALSE)
```

Arguments

`X`	is a data vector (numeric or ordered factor)
`W`	is a vector of weights
`withsqrt`	if TRUE function returns index given by BL2, elsewhere by BL (default). See more in details.

Details

Let m be the median category, n be the number of categories and P_i be the cumulative distribution of i-th category. The indices of Blair and Lacy (2000) are the following:

$$BL = 1 - \frac{\sum_{i=1}^{n-1}(P_i - 0.5)^2}{\frac{n-1}{4}}$$

$$BL2 = 1 - \left(\frac{\sum_{i=1}^{n-1}(P_i - 0.5)^2}{\frac{n-1}{4}}\right)^{\frac{1}{2}}$$

Value

The value of Blair and Lacy coefficient.

References

Blair J, Lacy M G. (2000): Statistics of ordinal variation, Sociological Methods and Research 28(251);251-280.

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
BL(X)
BL(X,W)

data(Well_being)
# Blair and Lacy index for health assessment with sample weights
X=Well_being$V1
W=Well_being$Weight
BL(X,W)
```

CoefVar *Coefficient of Variation*

Description

Computes Coefficient of Variation inequality measure of a given variable taking into account weights.

Usage

```
CoefVar(X, W = rep(1, length(X)), square = FALSE)
```

Arguments

X	is a data vector
W	is a vector of weights
square	logical, argument of the function CoefVar, for details see below

Details

Coefficient of variation is given by:

$$CV = \frac{\sigma}{\mu} \times 100$$

where σ is a standard deviation and μ is arithmetic mean.

Value

The value of CoefVar coefficient.

References

Sheret M.: (1984) Social Indicators Research, An International and Interdisciplinary Journal for Quality-of-Life Measurement, Vol. 15, No. 3, Oct. ISSN 03038300

Coulter P. B.: (1989) Measuring Inequality ISBN 0-8133-7726-9

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
CoefVar(X)
CoefVar(X,W)

data(Tourism)
#Coefficient of variation for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
CoefVar(X,W)
```

`Entropy` *Generalized entropy index*

Description

Computes generalized entropy index of a given variable taking into account weights.

Usage

```
Entropy(X, W = rep(1, length(X)), power = 0.5, zeroes = "include")
```

Arguments

`X`	is a data vector
`W`	is a vector of weights
`power`	is a entropy parameter
`zeroes`	defines what to do with zeroes in the data vector. Possible options are "remove" and "include". See Details for more.

Details

Entropy coefficient with respect to parameter α is equal to Theil_L(X,W) whenever $\alpha = 0$, is equal to Theil_T(X,W) whenever $\alpha = 1$, and whenever $\alpha \in (0, 1)$ we have

$$GE(\alpha) = \frac{1}{\alpha(\alpha-1)W}\sum_{i=1}^{n} w_i \left(\left(\frac{x_i}{\mu}\right)^{\alpha} - 1\right)$$

where W is a sum of weights and μ is the arithmetic mean of $x_1, ..., x_n$. Entropy coefficient is not well-defined for data vector with zero values whenever parameter is zero or one. In such case, entropy index coincides with the definition of Theil L index and Theil T index, respectively, and entropy index is calculated with corresponding Theil function. Theil L always removes zeroes. Theil T enables two ways to deal with zeroes by parameter zeroes. Option "remove" discard these X's and corresponding weights. Works for power>0. Option "include" puts $0 \log 0$ =0 due to limiting property of $p \log p$ in zero preserving zero value in dataset. It is valid only for Theil T index, that is power=0.

Value

The value of generalized entropy index

References

Shorrocks A. F.: (1980) The Class of Additively Decomposable Inequality Measures. Econometrica

Pielou E.C.: (1966) The measurement of diversity in different types of biological collections. Journal of Theoretical Biology

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Entropy(X)
Entropy(X,W)

data(Tourism)
# Generalized entropy index for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Entropy(X,W)
```

Gini *Gini coefficient*

Description

Computes Gini coefficient of a given variable taking into account weights.

Usage

```
Gini(X, W = rep(1, length(X)), fast = TRUE, rounded.weights = FALSE)
```

Arguments

`X`	is a data vector
`W`	is a vector of weights
`fast`	logical, if TRUE (default), Gini is calculated via matrix operations - fast but may cause memory allocation problems. If FALSE, Gini is calculated via vector operations - slower but with better memory allocation
`rounded.weights`	logical, may be run when fast=FALSE. If TRUE (default), Gini is calculated through alternative formula based on ordered X and integer weights. Choose it when dealing with memory allocation problems.

Details

Gini coefficient is given by:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mid x_i - x_j \mid}{2n^2\overline{x}}$$

Value

The value of Gini coefficient.

References

Dixon P. M., Weiner, J., Mitchell-Olds, T., and Woodley, R.: (1987) Bootstrapping the Gini Coefficient of Inequality. Ecology , Volume 68 (5)

Firebaugh G.: (1999) Empirics of World Income Inequality, American Journal of Sociology

Deininger K.; Squire L.: (1996) A New Data Set Measuring Income Inequality, The World Bank Economic Review, Vol. 10, No. 3

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Gini(X)
Gini(X,W)

data(Tourism)
#Gini coefficient for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Gini(X,W)
```

Hoover *Hoover index*

Description

Computes Hoover inequality measure of a given variable taking into account weights.

Usage

```
Hoover(X, W = rep(1, length(X)))
```

Arguments

X	is a data vector
W	is a vector of weights

Details

Let x_i be the income of the i-th person and $\overline{x}$ be the mean income. Then the Hoover index H is:

$$H = \frac{1}{2}\frac{\sum_i |x_i - \overline{x}|}{\sum_i x_i}$$

Value

The value of Hoover coefficient.

References

Hoover E. M. Jr.: (1936) The Measurement of Industrial Localization, The Review of Economics and Statistics, 18

Hoover E. M. Jr.: (1984) An Introduction to Regional Economics, ISBN 0-07-554440-7

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Hoover(X)
Hoover(X,W)

data(Tourism)
#Hoover index for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Hoover(X,W)
```

`ineq.weighted` *Weighted inequality measures*

Description

Calculates weighted mean and sum of X (or median of X), and a set of relevant inequality measures.

Usage

```
ineq.weighted(
  X,
  W = rep(1, length(X)),
  AF.norm = TRUE,
  Atkinson.e = 1,
  Jenkins.alfa = 0.8,
  Entropy.power = 0.5,
  zeroes = "include",
  Kolm.p = 1,
  Kolm.scale = "Standardization",
  Leti.norm = T,
  AN_Y.a = 1,
  AN_Y.b = 1,
  Apouey.a = 2/(1 - length(W[!is.na(W) & !is.na(X)])),
 Apouey.b = length(W[!is.na(W) & !is.na(X)])/(length(W[!is.na(W) & !is.na(X)]) - 1),
  BL.withsqrt = FALSE
)
```

Arguments

X	is a data vector
W	is a vector of weights
AF.norm	(logical). If TRUE (default) then index is divided by its maximum possible value
Atkinson.e	is a parameter for Atkinson coefficient
Jenkins.alfa	is a parameter for Jenkins coefficient
Entropy.power	is a generalized entropy index parameter
zeroes	defines what to do with zeroes in the data vector. Possible options are "remove" and "include". See Entropy function for details.
Kolm.p	is a parameter for Kolm index
Kolm.scale	method of data standardization before computing
Leti.norm	(logical). If TRUE (default) then Leti index is divided by a maximum possible value
AN_Y.a	is a positive parameter for Abul Naga and Yalcin inequality measure
AN_Y.b	is a parameter for Abul Naga and Yalcin inequality measure
Apouey.a	is a parameter for Apouey inequality measure
Apouey.b	is a parameter for Apouey inequality measure
BL.withsqrt	if TRUE function returns index given by BL2, elsewhere by BL (default). See more in details of BL function.

Details

Function checks if X is a numeric or an ordered factor. Then it calculates all appropriate inequality measures.

Value

The data frame with weighted mean and sum of X, and all inequality measures relevant for a numeric data. In a case of an ordered factor, the data frame with median of X, and all relevant inequality measures.

Examples

```
# Compare weighted and unweighted result.
X=1:10
W=1:10
ineq.weighted(X)
ineq.weighted(X,W)


data(Tourism)
# Results for Total expenditure with sample weights:
X=Tourism$`Total expenditure`
W=Tourism$`Sample weight`
ineq.weighted(X)
ineq.weighted(X,W)
```

`ineq.weighted.boot` *Weighted inequality measures with bootstrap*

Description

For weighted mean and weighted total of X (or median of X) as well as for each relevant inequality measure, returns outputs from ineq.weighted and bootstrap outcomes: expected value, bias (in %), standard deviation, coefficient of variation, lower and upper bound of confidence interval.

Usage

```
ineq.weighted.boot(
  X,
  W = rep(1, length(X)),
  B = 100,
  AF.norm = TRUE,
  Atkinson.e = 1,
  Jenkins.alfa = 0.8,
  Entropy.power = 0.5,
  zeroes = "include",
  Kolm.p = 1,
  Kolm.scale = "Standardization",
  Leti.norm = T,
  AN_Y.a = 1,
  AN_Y.b = 1,
  Apouey.a = 2/(1 - length(W[!is.na(W) & !is.na(X)])),
 Apouey.b = length(W[!is.na(W) & !is.na(X)])/(length(W[!is.na(W) & !is.na(X)]) - 1),
  BL.withsqrt = FALSE,
  keepSamples = FALSE,
  keepMeasures = FALSE,
  conf.alpha = 0.05,
  calib.boot = FALSE,
  Xs = rep(1, length(X)),
  total = sum(W),
  calib.method = "truncated",
  bounds = c(low = 0, upp = 10)
)
```

Arguments

`X`	is a data vector
`W`	is a vector of weights
`B`	is a number of bootstrap samples.
`AF.norm`	(logical). If TRUE (default) then index is divided by its maximum possible value
`Atkinson.e`	is a parameter for Atkinson coefficient

`Jenkins.alfa`	is a parameter for Jenkins coefficient
`Entropy.power`	is a generalized entropy index parameter
`zeroes`	defines what to do with zeroes in the data vector. Possible options are "remove" and "include". See Entropy function for details.
`Kolm.p`	is a parameter for Kolm index
`Kolm.scale`	method of data standardization before computing
`Leti.norm`	(logical). If TRUE (default) then Leti index is divided by a maximum possible value
`AN_Y.a`	is a positive parameter for Abul Naga and Yalcin inequality measure
`AN_Y.b`	is a parameter for Abul Naga and Yalcin inequality measure
`Apouey.a`	is a parameter for Apouey inequality measure
`Apouey.b`	is a parameter for Apouey inequality measure
`BL.withsqrt`	if TRUE function returns index given by BL2, elsewhere by BL (default). See more in details of BL function.
`keepSamples`	if TRUE, it returns bootstrap samples of data (Xb) and weights (Wb)
`keepMeasures`	if TRUE, it returns values of all inequality measures for each bootstrap sample
`conf.alpha`	significance level for confidence interval
`calib.boot`	if FALSE, then naive bootstrap is performed, calibrated bootstrap elsewhere
`Xs`	matrix of calibration variables. By default it is a vector of 1's, applied if calib.boot is TRUE
`total`	vector of population totals. By default it is a sum of weights, applied if calib.boot is TRUE
`calib.method`	weights' calibration method for function calib (sampling)
`bounds`	vector of bounds for the g-weights used in the truncated and logit methods; 'low' is the smallest value and 'upp' is the largest value

Details

By default, naive bootstrap is performed, that is no weights calibration is conducted. You can choose calibrated bootstrap to calibrate weights with respect to provided variables (Xs) and totals (total). Confidence interval is simply derived with quantile of order α and $1 - \alpha$ where α is a significance level for confidence interval.

Value

This functions returns a data frame from ineq.weighted extended with bootstrap results: expected value, bias (in %), standard deviation, coefficient of variation, lower and upper bound of confidence interval. If keepSamples=TRUE or keepMeasures==TRUE then the output becomes a list. If keepSamples=TRUE, the functions returns Xb and Wb, which are the samples of vector data and the samples of weights, respectively. If keepMeasures==TRUE, the functions returns Mb, which is a set of inequality measures from bootstrapping.

Examples

```
# Inequality measures with additional statistics for numeric variable
X=1:10
W=1:10
ineq.weighted.boot(X,W,B=10)

# Inequality measures with additional statistics for ordered factor variable
X=factor(c('H','H','M','M','L','L'),levels = c('L','M','H'),ordered = TRUE)
W=c(2,2,3,3,8,8)
ineq.weighted.boot(X,W,B=10)
```

Jenkins *Jenkins, Cowell and Flachaire*

Description

Computes Jenkins as well as Cowell and Flachaire inequality measure of a given variable taking into account weights.

Usage

```
Jenkins(X, W = rep(1, length(X)), alfa = 0.8)
```

Arguments

X	is a data vector
W	is a vector of weights
alfa	is the Jenkins coefficient parameter

Details

Jenkins coefficient is given by:

$$J = 1 - \sum_{j=0}^{K-1} (p_{j+1} - p_j)(GL_j + GL_{j+1})$$

where GL is Generalized Lorenz curve.

Cowell and Flachaire coefficient with alpha parameter is given by:

$$I(\alpha) = \frac{1}{\alpha(\alpha-1)}(\frac{1}{N}\sum_{i=1}^{N} s_i^{\alpha} - 1)$$

for $\alpha \in (0, 1)$, and

$$I(0) = -\frac{1}{N}\sum_{i=1}^{N} log(s_i)$$

for $\alpha = 0$.

Value

The value of Jenkins, Cowell and Flachaire coefficient.

References

Jenkins S. P. and P. J. Lambert: (1997) Three 'I's of Poverty Curves, with an Analysis of U.K. Poverty Trends

Cowell F. A.: (2000) Measurement of Inequality, Handbook of Income Distribution

Cowell F. A., Flachaire E.: (2017) Inequality with Ordinal Data

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Jenkins(X)
Jenkins(X,W)

data(Tourism)
#Jenkins, Cowell and Flachaire coefficients for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Jenkins(X,W)
```

`Kolm` *Kolm index*

Description

Computes Kolm inequality measure of a given variable taking into account weights.

Usage

```
Kolm(X, W = rep(1, length(X)), parameter = 1, scale = "None")
```

Arguments

`X`	is a data vector
`W`	is a vector of weights
`parameter`	is a Kolm parameter
`scale`	method of data scaling (None, Normalization, Unitarization, Standardization)

Details

Kolm index with parameter α is defined as:

$$K = \frac{1}{\alpha}(log(\sum_{i=1}^{n} \exp(\alpha(w_i - \mu)) - log(n)))$$

Kolm index is scale-dependent. Basic normalization methods can be applied before final computation.

Value

The value of Kolm coefficient.

References

Kolm S. C.: (1976) Unequal inequalities I and II

Kolm S. C.: (1996) Intermediate measures of inequality

Chakravarty S. R.: (2009) Inequality, Polarization and Poverty e-ISBN 978-0-387-79253-8

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Kolm(X)
Kolm(X,W)

# Compare raw and standardized data.
Kolm(X,W)
Kolm(X,W, scale ="Standardization")

# Changing units has an impact on the final result
Kolm(X)
Kolm(10*X)

# Changing units has no impact on the final result with standardized data
Kolm(X,scale ="Standardization")
Kolm(10*X,scale ="Standardization")
```

`Leti` *Leti index*

Description

Computes Leti inequality measure of a given variable taking into account weights.

Usage

```
Leti(X, W = rep(1, length(X)), norm = T)
```

Arguments

X	is a data vector (ordered factor or numeric)
W	is a vector of weights
norm	(logical). If TRUE (default) then Leti index is divided by a maximum possible value which is $(k-1)/2$ where k in a number of categories.

Details

Let n_i be the number of individuals in category i and let N be the total sample size. Cumulative distribution is given by $F_i = \frac{\sum_{j=1}^{i} n_j}{N}$. Leti index is defined as:

$$L = 2\sum_{i=1}^{k-1} F_i(1-F_i)$$

Value

The value of Leti coefficient.

References

Leti G.: (1983). Statistica descrittiva, il Mulino, Bologna. ISBN: 8-8150-0278-2

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Leti(X)
Leti(X,W)

data(Tourism)
#Leti index for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Leti(X,W)
```

LowerSum *Weighted lower sum*

Description

Computes weighted sum of values not greater then a quantile derived for the given probability.

Usage

```
LowerSum(X, W = rep(1, length(X)), p = 0.5)
```

Arguments

X	is a numeric data vector
W	is a vector of weights
p	is a probability to derive corresponding quantile

Details

Calculates weighted sum of values not greater then a quantile derived for the given probability based on cumulative distribution. Linear interpolation is applied to deal with a frequency distribution.

Value

The weighted sum of values not greater then a quantile.

Examples

```
# Suppose X represents incomes. Compare total incomes with incomes of poorer half of population.
X=1:10
W=10:1
sum(W*X)
LowerSum(X,W,0.5)
```

medianf *Median of ordered factor or numeric*

Description

Computes median of ordered factor or numeric variable taking into account weights.

Usage

```
medianf(X, W = rep(1, length(X)))
```

Arguments

X	is a data vector (numeric or ordered factor)
W	is a vector of weights

Details

Calculates median based on cumulative distribution. Tailored for ordered factors.

Value

The median category (number or label) of ordered factor.

Examples

```
# Compare weighted and unweighted result
X=factor(c('H','H','M','M','L','L'),levels = c('L','M','H'),ordered = TRUE)
W=c(2,2,3,3,8,8)
medianf(X)
medianf(X,W)
```

Palma *Palma index*

Description

Palma proportion - originally the ratio of the total income of the 10% richest people to the 40% poorest people.

Usage

```
Palma(X, W = rep(1, length(X)))
```

Arguments

X	is a data vector (numeric or ordered factor)
W	is a vector of weights

Details

Palma index is calculated by the following formula:

$$Palma = \frac{H}{L}$$

where H is share of 10% of the highest values, L is share of 40% of the lowest values.

Value

The value of Palma coefficient.

References

Cobham A., Sumner A.: (2013) Putting the Gini Back in the Bottle? 'The Palma' as a Policy-Relevant Measure of Inequality

Palma J. G.: (2011) Homogeneous middles vs. heterogeneous tails, and the end of the 'Inverted-U': the share of the rich is what it's all about

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Palma(X)
Palma(X,W)

data(Tourism)
#Palma index for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Palma(X,W)
```

Prop20_20 *Proportion 20:20*

Description

20:20 ratio - originally the ratio of the total income of the 20% richest people to the 20% poorest people.

Usage

```
Prop20_20(X, W = rep(1, length(X)))
```

Arguments

X	is a data vector (numeric or ordered factor)
W	is a vector of weights

Details

20:20 ratio is calculated as follows:

$$Prop = \frac{H}{L}$$

where H is share of 20% of the highest values, L is share of 20% of the lowest values.

Value

The value of 20:20 ratio coefficient.

References

Panel Data Econometrics: Theoretical Contributions And Empirical Applications edited by Badi Hani Baltag

Notes on Statistical Sources and Methods - The Equality Trust.

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Prop20_20(X)
Prop20_20(X,W)

data(Tourism)
#Prop20_20 proportion for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Prop20_20(X,W)
```

Quantile — *Sample quantile for weighted data*

Description

Computes quantile derived for the given probability taking into account weights.

Usage

```
Quantile(X, W = rep(1, length(X)), p = 0.5)
```

Arguments

`X`	is a numeric data vector
`W`	is a vector of weights
`p`	is a probability to derive corresponding quantile

Details

Linear interpolation is applied to deal with a frequency distribution.

Value

The quantile for weighted data.

Examples

```
# Compare weighted and unweighted result
X=1:10
W=10:1
Quantile(X,p=0.5)
Quantile(X,W,p=0.5)
```

RicciSchutz — *Ricci and Schutz index*

Description

Computes Ricci and Schutz inequality measure of a given variable taking into account weights.

Usage

```
RicciSchutz(X, W = rep(1, length(X)))
```

Arguments

X	is a data vector
W	is a vector of weights

Details

In the case of an empirical distribution with n elements where y_i denotes the wealth of household i and $\overline{y}$ the sample average, the Ricci and Schutz coefficient can be expressed as:

$$RS = \frac{1}{2n} \sum_{i=1}^{n} \frac{\mid y_i - \overline{y} \mid}{\overline{y}}$$

Value

The value of Ricci and Schutz coefficient.

References

Coulter P. B.: (1989) Measuring Inequality ISBN 0-8133-7726-9

Eliazar I. I., Sokolov I. M.: (2010) Measuring statistical heterogeneity: The Pietra index

Costa R. N., Pérez-Duarte S.: (2019) Not all inequality measures were created equal, Statistics Paper Series, No 31

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
RicciSchutz(X)
RicciSchutz(X,W)

data(Tourism)
#Ricci and Schutz index for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
RicciSchutz(X,W)
```

Theil_L *Theil L*

Description

Computes Theil_L inequality measure of a given variable taking into account weights.

Usage

```
Theil_L(X, W = rep(1, length(X)))
```

Arguments

X	is a data vector
W	is a vector of weights

Details

Theil L index is defined as:

$$T_L = T_{\alpha=0} = \frac{1}{N}\sum_{i=1}^{N} ln(\frac{\mu}{x_i})$$

where

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Theil L index can be computed only for positive values. By default, this functions discard zero X's and corresponding weights.

Value

The value of Theil_L coefficient.

References

Serebrenik A., van den Brand M.: Theil index for aggregation of software metrics values. 26th IEEE International Conference on Software Maintenance. IEEE Computer Society.

Conceição P., Ferreira P.: (2000) The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications

OECD: (2020) Regions and Cities at a Glance 2020, Chapter: Indexes and estimation techniques

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Theil_L(X)
Theil_L(X,W)

data(Tourism)
# Theil L coefficient for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Theil_L(X,W)
```

Theil_T *Theil T*

Description

Computes `Theil_T` inequality measure of a given variable taking into account weights.

Usage

```
Theil_T(X, W = rep(1, length(X)), zeroes = "include")
```

Arguments

`X`	is a data vector
`W`	is a vector of weights
`zeroes`	defines what to do with zeroes in the data vector. Possible options are "remove" and "include". See Details for more.

Details

Theil T index is defined as:

$$T_T = T_{\alpha=1} = \frac{1}{N}\sum_{i=1}^{N}\frac{x_i}{\mu}ln(\frac{x_i}{\mu})$$

where

$$\mu = \frac{1}{N}\sum_{i=1}^{N}x_i$$

Formally, Theil index is defined for positive values due to logarithms. Nevertheless, in data analysis zero values may occur. There are two way we can deal with them. Option "remove" discard these X's and corresponding weights. Option "include" puts $0\log 0$ =0 due to limiting property of $p\log p$ in zero preserving zero value in dataset.

Value

The value of `Theil_T` coefficient.

References

Serebrenik A., van den Brand M.: Theil index for aggregation of software metrics values. 26th IEEE International Conference on Software Maintenance. IEEE Computer Society.

Conceição P., Ferreira P.: (2000) The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications

OECD: (2020) Regions and Cities at a Glance 2020, Chapter: Indexes and estimation techniques

Examples

```
# Compare weighted and unweighted result
X=1:10
W=1:10
Theil_T(X)
Theil_T(X,W)

data(Tourism)
# Theil T coefficient for Total expenditure with sample weights
X=Tourism$Total_expenditure
W=Tourism$Sample_weight
Theil_T(X,W)
```

`Tourism` *Sample survey on trips*

Description

Data from sample survey on trips conducted in Polish households.

Usage

```
data(Tourism)
```

Format

A data frame with 5319 observations of 17 variables

- Year
- Country
- Country code
- World region
- Purpose of trip
- Accommodation type
- Number of trip's participants
- Nights spent
- Travel agency (organiser)
- Sample weight
- Total expenditure
- Expenditure for organiser
- Private expenditure
- Expenditure on accommodation
- Expenditure on restaurants & café
- Expenditure on transport
- Expenditure on commodities

Details

Answers were modified due to disclosure control. Data presents only part of full database.

`Well_being` *Sample survey on quality of life*

Description

Data from sample survey on quality of life conducted on Polish-Ukrainian border in 2015 and 2019.

Usage

```
data(Well_being)
```

Format

A data frame with 1197 observations of 27 variables

- Area. Rural and urban
- Gender. Male and female
- Year. Year of survey (2015 and 2019)
- V1. I have good opportunities to use my talents and skills at work
- V2. I am treated with respect by others at work
- V3. I have adequate opportunities for vacations or leisure activities
- V4. The quality of local services where (I) live is good
- V5. There is very little pollution from cars or other sources where I spend most of my time
- V6. There are parks and green areas near my residence
- V7. I have the freedom to plan my life the way I want to
- V8. I feel safe walking around my neighborhood during the day
- V9. Overall, to what extent are you currently satisfied with your life
- V10. Overall, to what extent do you feel that the things you do in life are worthwhile
- V11. How do you rate your health
- V12. How do you rate your work
- V13. How do you rate your sleep
- V14. How do you rate your leisure time
- V15. How do you rate your family life
- V16. How do you rate your community and public affairs life
- V17. How do you rate your personal plans
- V18. How do you rate your housing conditions
- V19. How do you rate your personal income
- V20. How do you rate your personal prospects
- V21. Does being part of the local community make you feel good about yourself
- V22. Do you have a say in what the local community is like
- V23. Is your neighborhood a good place for you to live
- Weight. Sample weight for each household

Details

Questions are on Likert scale: 1 - the worst assessment, 5 - the best assessment. Only 23 questions were selected out of over 100 questions. Answers were modified due to disclosure control.

Index

∗ **datasets**
    Tourism, 28
    Well_being, 29

AF, 2
AN_Y, 3
Apouey, 4
Atkinson, 5

BL, 6

CoefVar, 8

Entropy, 9

Gini, 10

Hoover, 11

ineq.weighted, 12
ineq.weighted.boot, 14

Jenkins, 16

Kolm, 17

Leti, 18
LowerSum, 20

medianf, 20

Palma, 21
Prop20_20, 22

Quantile, 23

RicciSchutz, 24

Theil_L, 25
Theil_T, 26
Tourism, 28

Well_being, 29