Package ‘ineqJD’

September 20, 2019

Type Package

Title Inequality Joint Decomposition

Version 1.0

Date 2019-09-03

Author Alberto Arcagni [aut, cre],
Igor Valli [aut],
Michele Zenga [ctb]

Maintainer Alberto Arcagni <alberto.arcagni@uniroma1.it>


License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2019-09-20 07:30:02 UTC

R topics documented:

bonferroni ................................................................. 2
dataProcessing ..................................................... 3
gini ................................................................. 4
inequalityCurves ................................................... 5
plot.inequality_curves ............................................ 7
summary.decomposition .......................................... 9
zenga ............................................................. 11

Index 13
bonferroni

Point and synthetic Bonferroni indexes

Description

Computes the decomposition of the Bonferroni point inequality indexes of a statistical variable \( Y \) described in the object \( x \).

Usage

bonferroni(x)

Arguments

\( x \)  
An object of class "dataProcessed". \( x \) is usually the result of \texttt{dataProcessing} function. More details are given in the "Details" section and \texttt{dataProcessing} help page.

Details

\texttt{bonferroni} computes the decomposition of the Bonferroni point inequality indexes from the object \( x \) of class "dataProcessed". \( x \) is usually the result of \texttt{dataProcessing} function.

Value

<table>
<thead>
<tr>
<th>index</th>
<th>String denoting computed index.</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{decomposition}</td>
<td>Array containing the decompositions. The dimensions of ( \texttt{decomposition} ) are ( c(g,g,r,s) ) where ( g ) is the number of groups, ( r ) the number of different values of ( Y ), and ( s ) the number of sources.</td>
</tr>
<tr>
<td>( x )</td>
<td>Object of class \texttt{dataProcessed} passed as input.</td>
</tr>
</tbody>
</table>

Author(s)

Alberto Arcagni, Igor Valli.

References


See Also

\texttt{gini} and \texttt{zenga} for other inequality indexes and \texttt{dataProcessing} for the class "dataProcessed".
dataProcessing

Examples

G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3) # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000) # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000, 600) # vector of the second source
data <- data.frame(G, X1, X2) # no sample weights are considered
x <- dataProcessing(# data preparation
  units = data[, c('X1', 'X2')],
  groups = data[, 'G'],
)
decomposition <- bonferroni(x)
decomposition

dataProcessing

Data Processing

Description

Convert raw data to frequency distribution framework and returns cumulative sums.

Usage

dataProcessing(
  units,
  groups = rep("G1", nrow(as.matrix(units)) ),
  weights = rep(1, nrow(as.matrix(units)))
)

Arguments

units Numeric vector of length n or matrix of dimension c(n,s) containing s sources referred to n statistical units

groups Vector of length n of group membership. If empty only one group is considered, otherwise the number of groups g is defined by the number of distinct values or levels in this vector.

weights Vector of lwights length n. If empty uniform weights are considered.

Details

dataProcessing convert raw data in the frequency distribution framework with r distinct values of Y. In this way repeated values are removed as well as ordering issues. Moreover cumulative frequencies and cumulative sources values are evaluated in order to prepare data for inequality decompositions.

Value

yh Vector of length r of distinct values of Y.

Phl Matrix of absolute cumulative frequencies of dimension c(r,g).

Qhlk Array of cumulative sum of sources of dimension c(r,g,s).


Author(s)

Alberto Arcagni, Igor Valli

References


Examples

G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3) # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000) # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000,600) # vector of the second source
data <- data.frame(G, X1, X2) # no sample weights are considered
x <- dataProcessing(
    units = data[, c('Var X1', 'Var X2')],
    groups = data[, 'G'],
)
x

gini

Point and synthetic Gini indexes

Description

Computes point and synthetic Gini indexes on a variable Y.

Usage

gini(x)

Arguments

x List containing: 'yh', the vector of unique values of the variable Y whose Bonferroni index is computed; 'Phl', the matrix of absolute cumulative frequencies; 'Qhlk', the matrix of cumulative sums of Y or its sources. x is usually the result of dataProcessing function. More details are given in the "Details" section and dataProcessing help page.

Details

gini compute point and synthetic Gini indexes on a variable y, e.g. income, on a statistical population that could be partitioned in g subpopulations and could be considered as sum of c sources, e.g. income sources.
inequalityCurves

Value

index   String denoting computed index.
decomposition array containing the decompositions.
x object usually of class dataProcessed passed as input.

Author(s)

Alberto Arcagni, Igor Valli.

References


Examples

G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3) # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000) # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000, 600) # vector of the second source
data <- data.frame(G, X1, X2) # no sample weights are considered
x <- dataProcessing( # data preparation
  units = data[, c('Var X1', 'Var X2')],
  groups = data[, 'Var G'],
)
decomposition <- gini(x)
decomposition

ingqualityCurves

Inequality curves evaluation

Description

Generates step-functions (see stepfun) representing inequality curves or sources/subpopulations point contributions in the decomposition x generated by functions gini, bonferroni or zenga.

Usage

inequalityCurves(x, ...)
## S3 method for class 'decomposition'
inequalityCurves(x, l = 1:dim(x$decomposition)[2], k = 1:dim(x$decomposition)[4], ...)
Arguments

- **x**: An object of class decomposition output of functions `gini`, `bonferroni`, `zenga`.
- **l**: Vector of selected subpopulations. If only one subpopulation is selected the resulting step function provides its point contributions. More than one subpopulation can be selected and contributions are cumulated. If empty all subpopulations are considered and contributions are marginalized by subpopulations.
- **k**: Vector of selected sources. If only one source is selected the resulting step function provides its point contributions. More than one source can be selected and contributions are cumulated. If empty all sources are considered and contributions are marginalized by sources.
- ...: Potentially further arguments (required by the generic).

Details

By default generates step functions representing inequality curves of the whole population. If arguments `l` and/or `k` are defined, results are step functions representing point contributions of the selected subpopulations and/or sources.

The class of the result is "inequality_curves" that is associated to the method `plot.inequality_curves` for graphical representations. Such derived class inerits the features of the "stepfun" class.

Value

An object of class "inequality_curves" inerits the features of class "stepfun" with the following additional attributes:

- **index**: String denoting computed index.
- **min, max**: The range of values assumed by the function.
- **groups**: Vector of names of the subpopulations partitioning the whole population.
- **sources**: Vector of names of the all the sources that sum to the total variable \( Y \).
- **selected_groups**: Vector of names of the selected subpopulations to evaluate the point contributions.
- **selected_sources**: Vector of names of the selected sources to evaluate the point contributions.

Author(s)

Alberto Arcagni, Igor Valli

References


See Also

See `gini`, `bonferroni`, `zenga` to obtain objects of class `decomposition` and see `plot.inequality_curves` for the graphical representation.

Examples

```r
G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3)  # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000)  # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000, 600)  # vector of the second source
data <- data.frame(G, X1, X2)  # no sample weights are considered
x <- dataProcessing(# data preparation
    units = data[, c("Var X1", "Var X2")],
    groups = data[, "Var G"],
)
decomposition <- zenga(x)

ic <- inequalityCurves(decomposition)
ic
contrib1 <- inequalityCurves(decomposition, l = 1)
contrib1
contrib12 <- inequalityCurves(decomposition, l = 1:2)
contrib12
```

Description

Method of the generic `plot` for objects generated by the function `inequalityCurves`.

Usage

```r
## S3 method for class 'inequality_curves'
plot(
x,
pch = 16,
from = 0,
to = 1,
xlim = NULL,
ylim = NULL,
xaxs = "i",
yaxs = "i",
xlab = "p",
ylab = NULL,
```
```
main = attributes(x)$index,
sub = paste0(
  "grp: ",
  paste(attributes(x)$groups, collapse = " ", ""); src: ",
  paste(attributes(x)$sources, collapse = " ", "
),
)
...  
```

**Arguments**

- **x**  
  Object of class "decomposition".

- **pch**  
  A vector of plotting characters or symbols: see `points`.

- **from, to**  
  The range over which the function will be plotted.

- **xlim**  
  The x limits (x1, x2) of the plot. The default value, NULL, indicates that the range of the function should be used.

- **ylim**  
  The y limits of the plot.

- **xaxs**  
  The style of axis interval calculation to be used for the x-axis. See `par` for details.

- **yaxs**  
  The style of axis interval calculation to be used for the y-axis. See `xaxs` above.

- **xlab**  
  A title for the x axis: see `title`.

- **ylab**  
  A title for the y axis: see `title`.

- **main**  
  An overall title for the plot: see `title`.

- **sub**  
  A sub title for the plot: see `title`.

- **...**  
  Arguments to be passed to methods, such as [graphical parameters](https://www.rdocumentation.org) (see `par`).

Many methods will accept the following arguments:

**Details**

This method is a convenience wrapper for plotting inequality curves. Default values of the `plot` are modified in order to plot inequality curves in the unitary square. Moreover, the default value of the argument `sub` shows in the plot if the curve represents the whole inequality or a contribution.

**Value**

The same output of the function `plot.stepfun`, a list with two components:

- **t**  
  Abscissa (x) values, including the two outermost ones.

- **y**  
  Y values ‘in between’ the t[].

**Author(s)**

Alberto Arcagni, Igor Valli
References


See Also

plot, graphical parameters, par, plot.stepfun, inequalityCurves

Examples

G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3) # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000) # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000, 600) # vector of the second source
data <- data.frame(G, X1, X2) # no sample weights are considered
x <- dataProcessing( # data preparation
  units = data[, c("Var X1", "Var X2")],
groups = data[, "G"],
)
decomposition <- zenga(x)
ic <- inequalityCurves(decomposition)
contrib1 <- inequalityCurves(decomposition, l = 1)
contrib12 <- inequalityCurves(decomposition, l = 1:2)
plot(ic)
plot(contrib1, add = TRUE)
plot(contrib12, add = TRUE)
text(0.1, 1/6+0:2/3, labels = c("G1", "G2", "G3"))

summary.decomposition  Summarizing inequality decomposition

Description

summary method for class "decomposition".

Usage

## S3 method for class 'decomposition'
summary(object, ...)
## S3 method for class 'summary.decomposition'
print(x, ...)
Arguments

object An object of class "decomposition", usually, as result of a call to \texttt{gini}, \texttt{bonferroni} and \texttt{zenga}.

x \ldots further arguments passed to or from other methods.

Details

\texttt{summary.decomposition} method use

Value

index String denoting computed index.

joint Array of joint decompositions by sources and subpopulations.

pairs Matrix of decompositions by subpopulations.

within Vector of within part to the overall inequality. It denotes the part of the overall inequality derived from the inequality inside each subpopulation.

between Vector of between part to the overall inequality. It denotes the part of the overall inequality derived from the comparison between subpopulations.

groups Vector of subpopulations contribution to the overall inequality.

groups_sources Matrix of subpopulations contributions for each source to the overall inequality.

sources Vector of sources contribution to the overall inequality.

synthetic Scalar denoting the value of the synthetic index.

Author(s)

Alberto Arcagni, Igor Valli.

References


See Also

\texttt{gini}, \texttt{bonferroni}, \texttt{zenga}, \texttt{dataProcessing}. 
Examples

G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3) # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000) # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000, 600) # vector of the second source
data <- data.frame(G, X1, X2) # no sample weights are considered
x <- dataProcessing( # data preparation
  units = data[, c('X1', 'X2')],
  groups = data[, 'G'],
)
decomposition <- zenga(x)

summary(decomposition)

zenga

Point and synthetic Zenga 2007 indexes

Description

Computes point and synthetic Zenga 2007 indexes on a variable \( Y \).

Usage

zenga(x)

Arguments

x

List containing: 'yh', the vector of unique values of the variable \( Y \) whose Bonferroni index is computed; 'Phl', the matrix of absolute cumulative frequencies; 'Qhlk', the matrix of cumulative sums of \( y \) or its sources. \( x \) is usually the result of dataProcessing function. More details are given in the "Details" section and dataProcessing help page.

Details

zenga compute point and synthetic Zenga 2007 indexes on a variable \( y \), e.g. income, on a statistical population that could be partitioned in \( g \) subpopulations and could be considered as sum of \( c \) sources, e.g. income sources.

Value

index String denoting computed index.
decomposition Array containing the decompositions.
x Object usually of class dataProcessed passed as input.

Author(s)

Alberto Arcagni, Igor Valli
References


Examples

G <- c(1, 2, 3, 1, 2, 3, 1, 1, 2, 3, 3, 3) # vector denoting group membership
X1 <- c(0, 0, 0, 500, 700, 300, 750, 1000, 500, 500, 500, 1000) # vector of the first source
X2 <- c(0, 0, 0, 500, 300, 700, 750, 500, 700, 700, 1000,600) # vector of the second source
data <- data.frame(G, X1, X2) # no sample weights are considered
x <- dataProcessing( # data preparation
  units = data[, c('Var X1', 'Var X2')],
  groups = data[, 'G'],
)

decomposition <- zenga(x)
decomposition
Index

*Topic Frequency Distribution
  dataProcessing, 3

bonferroni, 2, 5–7, 10

dataProcessing, 2, 3, 4, 10, 11

gini, 2, 4, 5–7, 10
  graphical parameters, 8, 9

inequalityCurves, 5, 7, 9

par, 8, 9

plot, 7, 9

plot.inequality_curves, 6, 7, 7

plot.stepfun, 8, 9

points, 8

print.dataProcessed (dataProcessing), 3
  print.summary.decomposition
    (summary.decomposition), 9

stepfun, 5, 6

summary.decomposition, 9

title, 8

zenga, 2, 5–7, 10, 11