Package 'ineq.2d'

July 22, 2025

Type Package

Title Two-Dimensional Decomposition of the Theil Index and the Squared
Coefficient of Variation

Version 0.1.0

Author Ivan Skliarov

Maintainer Ivan Skliarov <i.skliarov@student.uw.edu.pl>

Description Decomposition of income inequality by groups formed of individuals
possessing similar characteristics (e.g., sex, education, age) and their
income sources at the same time.
Decomposition of the Theil index is based on Giammatteo, M. (2007)
<https://www.lisdatacenter.org/wps/liswps/466.pdf>.
Decomposition of the squared coefficient of variation is based on
Garcia-Penalosa, C., & Orgiazzi, E. (2013) <doi:10.1111/roiw.12054>.

URL https://github.com/sklivan/ineq.2d

BugReports https://github.com/sklivan/ineq.2d/issues

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Depends R (>= 2.10)

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Repository CRAN

Date/Publication 2023-06-02 14:30:02 UTC

Contents

---

scv.2d	*Function performing two-dimensional decomposition of the squared coefficient of variation (SCV)*

Description

The function performs two-dimensional decomposition of the squared coefficient of variation according to Garcia-Penalosa & Orgiazzi (2013). That is, the coefficient can be decomposed by some feature that members of the studied population possess (e.g., sex, education, age) and their income source at the same time.

Usage

```
scv.2d(
  data,
  total,
  feature = NULL,
  sources = NULL,
  weights = NULL,
  perc = FALSE
)
```

Arguments

data	Data frame containing income data. Must contain at least one column with numeric values.
total	String specifying the name of the column containing data on total income.
feature	String specifying the name of the column containing information about the feature used for inequality decomposition. If left blank, total income is not decomposed by feature.
sources	Vector containing strings specifying the names of the columns with data on income sources, the sum of which must be equal to total income. If left blank, or the same value as in "total" is specified, then total income is not decomposed by income source.
weights	String specifying the name of the column containing population weights.
perc	If set to TRUE, then the function returns percentage shares of every inequality component in overall inequality. Set to FALSE by default.

Value

Data frame containing values of components of SCV.

Columns of the data frame represent values of the feature used for decomposition. There can be inequality within groups formed by this feature and between them - there are twice as much columns as values of the given feature. Whether a column contains a value of within or between-group inequality is indicated by ".W" and ".B" suffixes respectively.

Every row of the data frame represents an income source.

Thus, every value in this data frame is the contribution of inequality in income earned from i-th source by members of j-th population cohort to overall income inequality.

Remember that overall SCV, which is the sum of all values in the data frame, is always positive. However, some components of the coefficient can have negative contribution to inequality.

If all members of the studied population earn the same income, SCV normally must be equal to zero. But, scv.2d calculates "alpha," which is the absolute contribution of the given income source to overall inequality. Its calculation is impossible for identical incomes because the formula involves division by variance of income, which is zero in this case. Thus, the function will return the NaN value.

References

Garcia-Penalosa, C., & Orgiazzi, E. (2013). Factor Components of Inequality: A Cross-Country Study. Review of Income and Wealth, 59(4), 689-727.

Examples

```
# Load the test data set.
data("us16")

# No decomposition, just SCV of total income.
result <- scv.2d(us16, "hitotal", weights = "hpopwgt")

# Decomposition of income inequality by gender.
result <- scv.2d(us16, "hitotal", "sex", "hitotal", "hpopwgt")

# Decomposition of income inequality by gender and income source.
result <- scv.2d(us16, "hitotal", "sex", c("hilabour", "hicapital",
"hitransfer"), "hpopwgt")
```

---

theil.2d	*Function performing two-dimensional decomposition of the Theil index*

Description

The function performs two-dimensional decomposition of the Theil index according to Giammatteo (2007). That is, the index can be decomposed by some feature that members of the studied population possess (e.g., sex, education, age) and their income source at the same time.

The Theil index contains natural logarithm in its formula. This is why non-positive values of total income are removed during calculation.

Usage

```
theil.2d(
  data,
  total,
  feature = NULL,
  sources = NULL,
  weights = NULL,
  perc = FALSE
)
```

Arguments

data	Data frame containing income data. Must contain at least one column with numeric values.
total	String specifying the name of the column containing data on total income.
feature	String specifying the name of the column containing information about the feature used for inequality decomposition. If left blank, total income is not decomposed by feature.
sources	Vector containing strings specifying the names of the columns with data on income sources, the sum of which must be equal to total income. If left blank, or the same value as in "total" is specified, then total income is not decomposed by income source.
weights	String specifying the name of the column containing population weights.
perc	If set to TRUE, then the function returns percentage shares of every inequality component in overall inequality. Set to FALSE by default.

Value

Data frame containing values of components of the Theil index.

Columns of the data frame represent values of the feature used for decomposition. There can be inequality within groups formed by this feature and between them - there are twice as much columns as values of the given feature. Whether a column contains a value of within or between-group inequality is indicated by ".W" and ".B" suffixes respectively.

Every row of the data frame represents an income source.

Thus, every value in this data frame is the contribution of inequality in income earned from i-th source by members of j-th population cohort to overall income inequality.

Remember that overall Theil index, which is the sum of all values in the data frame, is always positive. However, some components of the index can have negative contribution to inequality.

References

Giammatteo, M. (2007). The Bidimensional Decomposition of Inequality: A nested Theil Approach. LIS Working papers, Article 466, 1-30.

Examples

```
# Load the test data set.
data("us16")

# No decomposition, just Theil index of total income.
result <- theil.2d(us16, "hitotal", weights = "hpopwgt")

# Decomposition of income inequality by gender.
result <- theil.2d(us16, "hitotal", "sex", "hitotal", "hpopwgt")

# Decomposition of income inequality by gender and income source.
result <- theil.2d(us16, "hitotal", "sex", c("hilabour", "hicapital",
"hitransfer"), "hpopwgt")
```

---

us16   *Test dataset for the ineq.2d package*

Description

This data set is a combination of household and personal-level data sets available as sample files on the Luxembourg Income Study website. The data sets were combined to assign personal characteristics of household heads to households.

Usage

```
us16
```

Format

A data frame with 1000 rows and 8 variables:

hitotal Total income, sum of hitransfer, hilabour, and hicapital.

hitransfer Transfer income.

hilabour Labor income.

hicapital Capital income.

hpopwgt Population weights.

age Age of the household head.

sex Sex of the household head.

educ Education level of the household head.

Details

Thus, the combined data set contains income data on 1000 households. This includes three income sources and value for total income. Additionally, contains several features of every household head: sex, education, and age.

Source

LIS Cross-national Data Center, https://www.lisdatacenter.org/resources/self-teaching/

Examples

```
# To load the dataset to the environment, use the following code:
data(us16)
```

Index