Package ‘mded’

April 27, 2015

Type Package
Title Measuring the Difference Between Two Empirical Distributions
Version 0.1-2
Date 2015-04-27
Author Hideo Aizaki
Maintainer Hideo Aizaki <azk-r@spa.nifty.com>
Description Provides a function for measuring the difference between two independent or non-independent empirical distributions and returning a significance level of the difference.
License CC0

R topics documented:

mded-package .......................................................... 1
mded ................................................................. 2

Index

mded-package      Measuring the difference between two empirical distributions

Description

The package provides a function for measuring the difference between two independent or non-independent empirical distributions and returning a significance level of the difference.

Acknowledgments

I would like to thank Professor Gregory L. Poe for his kindness.

1
Note

Recommended citations:


Author(s)

Hideo Aizaki

References


---

**mded**

*Measuring the difference between two empirical distributions*

**Description**

The function measures the difference between two independent or non-independent empirical distributions and returns a significance level of the difference.

**Usage**

```r
demed(distr1, distr2, detail = FALSE, independent = TRUE)
```

```r
## S3 method for class 'mded'
print(x, digits = max(3, getOption("digits") - 3), ...)```
**Arguments**

- **distr1**: A vector of empirical distribution. `distr1` is greater than `distr2`.
- **distr2**: A vector of empirical distribution.
- **detail**: If `TRUE`, a vector of the difference between `distr1` and `distr2` is returned.
- **independent**: Set as `FALSE` when `distr1` and `distr2` are not independent of each other.
- **x**: An object of S3 class 'mded.'
- **digits**: A number of significant digits.
- **...**: Arguments passed to the function print.

**Details**

The function measures the difference between two independent or non-independent empirical distributions and returns a significance level of the difference on the basis of the methods proposed by Poe et al. (1997, 2005). Such calculations are frequently needed in empirical econometric studies wherein (marginal) willingness-to-pay distributions that are estimated using contingent valuation methods or discrete choice experiments have to be compared to each other.

Let us assume that X and Y are empirical distributions, which are depicted by the vector \( x = (x_1, x_2, ..., x_m) \), and \( y = (y_1, y_2, ..., y_n) \). The null hypothesis (H0) is \( X - Y = 0 \), while the alternative hypothesis (H1) is \( X - Y > 0 \). When X and Y are independent of each other, the complete combinatorial method (Poe et al. 2005) provides the one-sided significance level of H0 that is calculated by \( \frac{\# \{xi - yj <= 0\}}{m * n} \), where \#\{cond\} provides the number of times that `cond` is true. When X and Y are not independent of each other, the pair difference method (Poe et al. 1997) provides the one-sided significance level of H0 that is calculated by \( \frac{\# \{xi - yi <= 0\}}{m} \), where \( m = n \).

Note that the function may take quite long, and would require large amount of memory to calculate the difference between two independent distributions if the argument `detail` is set as `TRUE` because the resulting difference is stored as a vector. For example, when `distr1` and `distr2` each contain 10,000 elements (observations), the vector of the difference contains 100,000,000 elements. If memory is lacking, R would stop running the function, showing an error message related to memory limitation.

**Value**

- **stat**: One-side significance level of the difference between `distr1` and `distr2`.
- **means**: A vector of mean values of `distr1` and `distr2`.
- **cases**: A vector of integer values describing a number of cases wherein the `cond` is true and that is false.
- **distr1**: A vector assigned to `distr1`.
- **distr2**: A vector assigned to `distr2`.
- **distr.names**: A vector of the names of objects assigned to `distr1` and `distr2`.
- **diff**: A vector of the difference. If `detail = TRUE`, it is returned.

**Author(s)**

Hideo Aizaki
References


Examples

```r
set.seed(123)
x <- rnorm(100, 3)
y <- rnorm(100, 1)

out <- mded(distr1 = x, distr2 = y, detail = TRUE)
out
```
Index

*Topic htest
  mded, 2

*Topic package
  mded-package, 1

mded, 2
mded-package, 1

print.mded (mded), 2