Package ‘DECIDE’

February 19, 2015

Type Package

Title DEComposition of Indirect and Direct Effects

Version 1.2

Date 27-11-2014

Author Christiana Kartsonaki

Maintainer Christiana Kartsonaki <christiana.kartsonaki@gmail.com>

Description Calculates various estimates for measures of educational differentials, the relative importance of primary and secondary effects in the creation of such differentials and compares the estimates obtained from two datasets.

License GPL (>= 2)

LazyLoad yes

Repository CRAN

Date/Publication 2014-11-27 16:04:40

NeedsCompilation no

R topics documented:

DECIDE-package .................................................. 2
compare.relimp .................................................. 2
create.classdata ............................................... 4
plot_transition ................................................ 5
prepare.data .................................................. 6
print_relimp .................................................. 6
relative.importance .......................................... 7

Index 10
compare.relimp

### Description

Calculates various estimates for measures of educational differentials, the relative importance of primary and secondary effects in the creation of such differentials and compares the estimates obtained from two datasets.

### Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>DECIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>1.2</td>
</tr>
<tr>
<td>Date:</td>
<td>2014-11-24</td>
</tr>
<tr>
<td>License:</td>
<td>GPL (&gt;= 2)</td>
</tr>
<tr>
<td>LazyLoad:</td>
<td>yes</td>
</tr>
</tbody>
</table>

See `relativeNimportance`.

### Author(s)

Christiana Kartsonaki  
Maintainer: Christiana Kartsonaki <christiana.kartsonaki@gmail.com>

### References


---

**compare.relimp**  
**Compare estimates of log odds, log odds ratios and relative importance obtained by two datasets**

### Description

Computes 95% confidence intervals for the differences in log odds of transition, log odds ratios and relative importance estimates between the two datasets. Also calculates chi-squared test statistics and p-values for testing whether the differences are different from zero.
Usage

```r
compare.relimp(dataset1, dataset2)
```

Arguments

- `dataset1` is the first dataset; a data frame with 4 columns, in the following order: 1: student’s ID, 2: class, 3: transition (0 if not, 1 if yes) and 4: performance score.
- `dataset2` is the second dataset; a data frame with 4 columns, in the following order: 1: student’s ID, 2: class, 3: transition (0 if not, 1 if yes) and 4: performance score.

Value

- `ci.diff.lo` 95% confidence intervals for differences in log odds of transition
- `test.diff.lo` Test statistic for differences in log odds
- `test.diff.lo.pvalue` p-value for testing for differences in log odds
- `ci.diff.lo.r` 95% confidence intervals for differences in log odds ratios
- `test.diff.lo` Test statistic for differences in log odds ratios
- `test.diff.lo.pvalue` p-value for testing for differences in log odds ratios
- `ci.diff.ri.1` 95% confidence intervals for relative importance estimates - 1
- `ci.diff.ri.2` 95% confidence intervals for relative importance estimates - 2
- `ci.diff.ri.avg` 95% confidence intervals for relative importance estimates - average

Author(s)

Christiana Kartsonaki

References


Examples

```r
# generate two datasets
set.seed(1)
data1 <- data.frame(seq(1:10), rep(c(1, 2), length.out = 10),
c(rep(0, times = 3), rep(1, times = 7)),
c(rnorm(4, 0, 1), rnorm(4, 0.5, 1), NA, NA))
```
create.classdata

Create data frames for each class

Description
Takes a data frame and creates a list of data frames by splitting the data by the factor "class".

Usage
create.classdata(dataset)

Arguments
dataset A data frame produced by prepare.dataset.

Value
data_class A list with number of elements equal to the number of classes and each element a data frame for each class.

Author(s)
Christiana Kartsonaki

Examples
# generate a dataset
data <- data.frame(seq(1:10), rep(c(1, 2), length.out = 10),
c(rep(0, times = 5), rep(1, times = 5)),
c(rnorm(5, 1, 1), rnorm(5, 0.5, 1)))

# run function
compare.relimp(data1, data2)

data2 <- data.frame(seq(1:10), rep(c(1, 2), length.out = 10),
c(rep(0, times = 5), rep(1, times = 5)),
c(rnorm(5, 1, 1), rnorm(5, 0.5, 1)))

# run function
compare.relimp(data1, data2)
**plot_transition**

**Plot distributions of performance and transition propensities**

**Description**
Plots distribution of academic performance and probabilities of transition for each class.

**Usage**

```
plot_transition(dataset)
```

**Arguments**

- `dataset` A data frame with 4 columns only, in the following order: 1: student’s ID, 2: class, 3: transition (0 if not, 1 if yes) and 4: performance score.

**Value**
A plot of the distributions of performance and transition propensities for each class.

**Author(s)**
Christiana Kartsonaki

**References**


**Examples**

```r
# generate a dataset
set.seed(1)
data <- data.frame(seq(1:10), rep(c(1, 2), length.out = 10),
c(rep(0, times = 3), rep(1, times = 7)),
c(rnorm(4, 0, 1), rnorm(4, 0.5, 1), NA, NA))

# run function
plot_transition(data)
```
### prepare.data

**Prepare dataset to be used in relative.importance**

**Description**

Prepares datasets to be in the format required by the function `relative.importance`. It is automatically called by `relative.importance`.

**Usage**

```r
prepare.data(dataset)
```

**Arguments**

- **dataset**: A data frame with 4 columns only, in the following order: 1: student’s ID, 2: class, 3: transition (0 if not, 1 if yes) and 4: performance score.

**Value**

- **dataset**: The data frame given as the argument, with column names changed and missing values removed.

**Author(s)**

Christiana Kartsonaki

**Examples**

```r
# generate a dataset
data <- data.frame(seq(1:10), rep(c(1, 2, 3, length.out = 10), rbinom(1, n = 10, p = 0.7), c(rnorm(8, 0, 1), NA, NA))

# run function
data_clean <- prepare.data(data)
```

### print_relimp

**Print tables of estimates**

**Description**

Presents various estimates for measures of educational differentials, the relative importance of primary and secondary effects and corresponding standard errors and confidence intervals.

**Usage**

```r
print_relimp(dataset)
```
relative.importance

Arguments

    dataset  A data frame with 4 columns only, in the following order: 1: student’s ID, 2: class, 3: transition (0 if not, 1 if yes) and 4: performance score.

Value

    Returns a more nicely presented version of the results given by relative.importance.

Author(s)

    Christiana Kartsonaki

References


See Also

    relative.importance

Examples

    # generate a dataset
    set.seed(1)
    data <- data.frame(seq(1:10), rep(c(1, 2, 3), length.out = 10), rbinom(1, n = 10, p = 0.7), c(rnorm(8, 0, 1), NA, NA))
    # run function
    print_relimp(data)

relative.importance  Relative importance of primary and secondary effects

Description

    Calculates various estimates for measures of educational differentials, the relative importance of primary and secondary effects and corresponding standard errors and confidence intervals.

Usage

    relative.importance(dataset)
relative.importance

Arguments

dataset A data frame with 4 columns only, in the following order: 1: student’s ID, 2: class, 3: transition (0 if not, 1 if yes) and 4: performance score.

Value

sample_size Total number of individuals
no_classes Number of classes
class_size A list of no_classes elements, each element containing the size of each class
percentage_overall Overall percentage that made the transition
percentage_class A list of no_classes elements, each element containing percentage that made the transition for each class
fifty_point 50% point of transition
parameters A data frame with the parameters of logistic regression ($\alpha, \beta$) and normal distribution ($\mu, \sigma$) for each class
transition_prob A data frame with the transition probabilities
log_odds A data frame with log odds of transition (diagonal elements: actual log odds for each class, off-diagonal: counterfactual log odds)
se_logodds A data frame with the standard errors of the log odds of transition
ci_logodds Approximate 95% confidence intervals for the log odds of transition
odds Odds of transition
log_oddsratios Log odds ratios
se_logoddsratios Standard errors for the log odds ratios
ci_logoddsratios Approximate 95% confidence intervals for the log odds ratios
oddsratios Odds ratios
rel_imp_prim1 Estimates of the relative importance of primary effects using the first equation for calculating the relative importance
rel_imp_prim2 Estimates of the relative importance of primary effects using the second equation for calculating the relative importance
rel_imp_prim_avg Estimates of the relative importance of primary effects using the the average of the two equations for calculating the relative importance
rel_imp_sec1 Estimates of the relative importance of secondary effects using the first equation for calculating the relative importance
rel_imp_sec2 Estimates of the relative importance of secondary effects using the second equation for calculating the relative importance
Estimates of the relative importance of secondary effects using the average of the two equations for calculating the relative importance.

**se.ri.1** Standard errors of the relative importance estimates given by the first equation.

**ci.ri.1** Approximate 95% confidence intervals for the relative importance of secondary effects given by the first equation.

**se.ri.2** Standard errors of the relative importance estimates given by the second equation.

**ci.ri.2** Approximate 95% confidence intervals for the relative importance of secondary effects given by the second equation.

**se.ri.avg** Standard errors of the relative importance estimates given by the average of the two equations.

**ci.ri.avg** Approximate 95% confidence intervals for the relative importance of secondary effects given by the average of the two equations.

**Author(s)**

Christiana Kartsonaki

**References**


**See Also**

`print_relimp`, `plot_transition`

**Examples**

```r
# generate a dataset
set.seed(1)
data <- data.frame(seq(1:10), rep(c(1, 2), length.out = 10),
c(rep(0, times = 3), rep(1, times = 7)),
c(rnorm(4, 0, 1), rnorm(4, 0.5, 1), NA, NA))

# run function
relative.importance(data)
```
Index

*Topic package
  DECIDE-package, 2

compare.relimp, 2
create.classdata, 4

DECIDE (DECIDE-package), 2
DECIDE-package, 2

plot_transition, 5, 9
prepare.data, 6
print.relimp, 6, 9

relative.importance, 7, 7