Package ‘cacc’

October 28, 2022

Title Conjunctive Analysis of Case Configurations

Version 0.1.0

Date 2022-10-20

Description A set of functions to conduct Conjunctive Analysis of Case Configurations (CACC) as described in Miethe, Hart, and Regoecri (2008) <doi:10.1007/s10940-008-9044-8>, and identify and quantify situational clustering in dominant case configurations as described in Hart (2019) <doi:10.1177/0011128719866123>. Initially conceived as an exploratory technique for multivariate analysis of categorical data, CACC has developed to include formal statistical tests that can be applied in a wide variety of contexts. This technique allows examining composite profiles of different units of analysis in an alternative way to variable-oriented methods.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.1

URL https://github.com/amoneva/cacc

BugReports https://github.com/amoneva/cacc/issues

Imports dplyr, ggplot2, rlang, stats, tibble, tidyr

Depends R (>= 2.10)

LazyData true

NeedsCompilation no

Author Asier Moneva [aut, cre] (<https://orcid.org/0000-0002-2156-0213>), Miriam Esteve [aut], Timothy C. Hart [aut]

Maintainer Asier Moneva <amoneva@nsr.nl>

Repository CRAN

Date/Publication 2022-10-28 17:10:12 UTC
R topics documented:

```
cacc ................................................................. 2
cluster_sci .......................................................... 3
cluster_xsq ........................................................... 4
main_effect ............................................................ 4
onharassment ............................................................ 5
plot_effect ............................................................... 6
plot_sci ................................................................. 7
```

Index 8

<table>
<thead>
<tr>
<th>cacc</th>
<th>Conjunctive Analysis of Case Configurations</th>
</tr>
</thead>
</table>

Description

Computes a Conjunctive Analysis of Case Configurations (CACC).

Usage

```
cacc(data, ivs, dv)
```

Arguments

```
data A data frame or a tibble.
ivs A vector of names of the independent variables, without quotes. Variables must be categorical, either integer, character, or factor.
dv Name of the dependent variable, without quotes. Variable must be a dichotomous integer or factor with values 0 (absence) and 1 (presence).
```

Value

Returns a tibble with the CACC matrix.

References

### Examples

```r
cacc(
  data = onharassment,
  ivs = c(sex, age, hours, snapchat, instagram, facebook, twitter, name, photos, privacy),
  dv = rep_victim
)
cacc(onharassment, ivs = sex:privacy, dv = rep_victim)

# Syntax with the native R pipe
onharassment |> cacc(ivs = sex:privacy, dv = rep_victim)
```

---

<table>
<thead>
<tr>
<th>cluster_sci</th>
<th>Situational Clustering Index</th>
</tr>
</thead>
</table>

### Description

Computes a Situational Clustering Index (SCI) to quantify the magnitude of the clustering of observations among dominant profiles in a `cacc_matrix`.

### Usage

```r
cluster_sci(cacc_matrix)
```

### Arguments

- **cacc_matrix**  
  A tibble. The output of the `cacc` function.

### Value

Returns a numeric value.

### References


### Examples

```r
cluster_sci(cacc(onharassment, ivs = sex:privacy, dv = rep_victim))
```
cluster_xsq  
*Chi-Square Goodness-of-Fit Test*

**Description**
Computes a Chi-Square Goodness-of-Fit Test to determine whether there is statistically significant clustering of observations among dominant profiles in a `cacc_matrix`.

**Usage**
```r
cluster_xsq(cacc_matrix)
```

**Arguments**
- `cacc_matrix`  
  A tibble. The output of the `cacc` function.

**Value**
Returns a list with the Chi-square results. This is the same object returned by the `chisq.test` function from the `stats` package.

**References**

**Examples**
```r
cluster_xsq(cacc(onharassment, ivs = sex:privacy, dv = rep_victim))
```

---

main_effect  
*Main effect*

**Description**
Computes the main effect that a specific value of a variable produces on the outcome probability in a `cacc_matrix`.

**Usage**
```r
main_effect(cacc_matrix, iv, value, summary = TRUE)
```

---
Arguments

- `cacc_matrix` A tibble. The output of the `cacc` function.
- `iv` A single variable name contained in a `cacc_matrix`.
- `value` A single numeric or character value the `iv` specified can take.
- `summary` Logical. Defaults to `TRUE`. Whether or not to return the summary statistics for the main effect.

Value

When `summary = TRUE`, returns a tibble with summary stats for the main effect. If `summary = FALSE`, returns a tibble containing a single numeric variable, ranging from 0 to 1, containing the main effects of the value of the selected `iv` on the probability of outcome.

References


Examples

```r
main_effect(
  cacc_matrix = cacc(onharassment, ivs = sex:privacy, dv = rep_victim),
  iv = age,
  value = "15-17"
)

main_effect(
  cacc_matrix = cacc(onharassment, ivs = sex:privacy, dv = rep_victim),
  iv = age,
  value = "15-17",
  summary = FALSE
)
```

onharassment  Profiles of 4174 Spanish students

Description

A dataset containing the demographics, online routines, and self-reported online harassment repeat victimization and offending of 4174 Spanish non-university education students.

Usage

```r
onharassment
```
Format

A data frame with 4174 rows and 12 variables:

- **sex** Factor. The students’ self-reported sex.
- **age** Factor. The students’ self-reported age.
- **hours** Factor. The students’ self-reported number of daily hours spent online.
- **snapchat** Factor. Whether students report using the social media Snapchat on a daily basis.
- **instagram** Factor. Whether students report using the social media Instagram on a daily basis.
- **facebook** Factor. Whether students report using the social media Facebook on a daily basis.
- **twitter** Factor. Whether students report using the social media Twitter on a daily basis.
- **name** Factor. Whether students report using their real names on social media.
- **photos** Factor. Whether students report uploading personal photos to social media.
- **privacy** Factor. Whether students report restricting their social media access to contacts only.
- **rep_victim** Factor. Whether students report repeatedly suffering online harassment.
- **rep_offender** Factor. Whether students report repeatedly committing online harassment.

Source


---

**plot_effect**

*Density Plot for the Main Effect*

Description

Plots an annotated boxplot and kernel density estimate to visualize the distribution of the main effect that a specific value of a variable produces on the outcome probability in a `cacc_matrix`.

Usage

`plot_effect(cacc_matrix, iv, value)`

Arguments

- **cacc_matrix** A tibble. The output of the `cacc` function.
- **iv** A single variable name contained in a `cacc_matrix`.
- **value** A single numeric or character value the `iv` specified can take.

Value

Returns a ggplot object.
References


Examples

```r
plot_effect(
  cacc_matrix = cacc(onharassment, ivs = sex:privacy, dv = rep_victim),
  iv = age,
  value = "15-17"
)
```

plot_sci

*Lorenz Curve for the Situational Clustering Index*

Description

Plots a Lorenz Curve for the Situational Clustering Index (SCI) to visualize the magnitude of the clustering of observations among dominant profiles in a `cacc_matrix`.

Usage

```r
plot_sci(cacc_matrix)
```

Arguments

- `cacc_matrix`: A tibble. The output of the `cacc` function.

Value

Returns a ggplot object.

References


Examples

```r
plot_sci(cacc_matrix = cacc(onharassment, ivs = sex:privacy, dv = rep_victim))
```
Index

* datasets
  onharassment, 5
  cacc, 2
  cluster_sci, 3
  cluster_xsq, 4
  main_effect, 4
  onharassment, 5
  plot_effect, 6
  plot_sci, 7