Package ‘wec’

October 31, 2017

Type Package

Title Weighted Effect Coding

Version 0.4-1

Date 2017-10-30

Author Rense Nieuwenhuis, Manfred te Grotenhuis, Ben Pelzer, Alexander Schmidt, Ruben Konig, Rob Eisinga

Maintainer Rense Nieuwenhuis <rense.nieuwenhuis@sofi.su.se>

Description Provides functions to create factor variables with contrasts based on weighted effect coding, and their interactions. In weighted effect coding the estimates from a first order regression model show the deviations per group from the sample mean. This is especially useful when a researcher has no directional hypotheses and uses a sample from a population in which the number of observation per group is different.

License GPL-3

Imports dplyr

URL http://www.ru.nl/sociology/mt/wec/downloads/

NeedsCompilation no

Repository CRAN

Date/Publication 2017-10-31 14:34:46 UTC

R topics documented:

BMI ................................................................. 2
contr.wec ...................................................... 3
PUMS ............................................................ 4
wec.interact .................................................. 5

Index 8
Description

The BMI data contains information on Dutch individuals’ BMI, in addition to select socio-demographic variables.

Format

A data frame with 3323 observations on the following 6 variables.

- `sex` a factor with levels `male` and `female`
- `education` a factor with levels `lowest`, `middle`, and `highest`
- `year` a factor with levels `2000`, `2005`, and `2011`
- `BMI` interval variable representing respondents’ Body Mass Index (BMI)
- `childless` a factor with levels `no` and `yes`
- `log_age` interval variable representing the natural log of respondents’ age
- `age_categorical` a factor with levels `Young (18-30)`, `Middle (31-59)` and `Older (60-70)`

Source

These data are a subset from three waves of the ‘Socio-Cultural Developments in the Netherlands’ (SOCON) datasets, collected at the Radboud University in the Netherlands (see references for original codebooks).

References


Examples

```r
data(BMI)

# Without Controls
model.dummy <- lm(BMI ~ education, data=BMI)
```
contr.wec

summary(model.dummy)

# With Controls
model.dummy.controls <- lm(BMI ~ education + sex + log_age + childless + year, data=BMI)
summary(model.dummy.controls)

---

**contr.wec**

*Function calculates contrasts for a factor variable based on weighted effect coding.*

**Description**

This function calculates contrasts for a factor variable based on weighted effect coding. In weighted effect coding, the estimates from a first order regression model show the deviations per group from the sample mean. This is especially useful when a researcher has no directional hypotheses and uses a sample from a population in which the number of observations per group is different.

**Usage**

```r
contr.wec(x, omitted)
```

**Arguments**

- `x` Factor variable
- `omitted` Label of the factor label that should be taken as the omitted category

**Value**

Returns a contrast matrix based on weighted effect coding.

**Author(s)**

Rense Nieuwenhuis, Manfred te Grotenhuis, Ben Pelzer, Alexander Schmidt, Ruben Konig, Rob Eisinga

**References**


See Also

wec.interact

Examples

data(BMI)

# Without controls
BMI$educ.wec.lowest <- BMI$educ.wec.highest <- BMI$educ

contrasts(BMI$educ.wec.lowest) <- contr.wec(BMI$education, omitted="lowest")
contrasts(BMI$educ.wec.highest) <- contr.wec(BMI$education, omitted="highest")

model.wec.lowest <- lm(BMI ~ educ.wec.lowest, data=BMI)
model.wec.highest <- lm(BMI ~ educ.wec.highest, data=BMI)
summary(model.wec.lowest)
summary(model.wec.highest)

# With Controls
BMI$sex.wec.female <- BMI$sex.wec.male <- BMI$sex
contrasts(BMI$sex.wec.female) <- contr.wec(BMI$sex, omitted="female")
contrasts(BMI$sex.wec.male) <- contr.wec(BMI$sex, omitted="male")

BMI$year.wec.2000 <- BMI$year.wec.2011 <- BMI$year
contrasts(BMI$year.wec.2000) <- contr.wec(BMI$year, omitted="2000")
contrasts(BMI$year.wec.2011) <- contr.wec(BMI$year, omitted="2011")

model.wec.lowest.controls <- lm(BMI ~ educ.wec.lowest +
       sex.wec.female + log_age + year.wec.2000,
       data=BMI)
model.wec.highest.controls <- lm(BMI ~ educ.wec.higher +
       sex.wec.male + log_age + year.wec.2011,
       data=BMI)
summary(model.wec.lowest.controls)
summary(model.wec.highest.controls)

PUMS  Public Use Microdata Sample files (PUMS) 2013

Description

The ACS Public Use Microdata Sample files (PUMS) are a sample of the actual responses to the American Community Survey and include most population and housing characteristics.
**wec.interact**

**Format**

A data frame with 10000 observations on the following 4 variables.

- **wage**: annual wages (binned to 1000s, top-coded, in US dollar)
- **race**: a factor with levels Hispanic, Black, Asian, and White
- **education.int**: level of education
- **education.cat**: a factor variable with levels High School, and Degree

**Source**

These data are a random subset of 10000 observations from working individuals aged over 25 in the 2013 ACS Public Use Microdata Sample files (PUMS).

**Examples**

```r
data(PUMS)
PUMS$race.wec <- factor(PUMS$race)
contrasts(PUMS$race.wec) <- contr.wec(PUMS$race.wec, "White")

m.wec <- lm(wage ~ race.wec, data=PUMS)
summary(m.wec)

PUMS$race.educint <- wec.interact(PUMS$race.wec, PUMS$education.int)
m.wec.educ <- lm(wage ~ race.wec + education.int + race.educint, data=PUMS)
summary(m.wec.educ)
```

---

**wec.interact**  
*Function to create an interaction between two variables based on weighted effect coding.*

**Description**

This function facilitates the estimation of an interaction between two factor variables that are based on weighted effect coding. To that end, it creates a third variable that, together with the two original factor variables, forms the complete interaction. In interaction models, weighted effect coding displays the extra effect on top of the main effects found in a model without the interaction effect(s).

**Usage**

```r
wec.interact(x1, x2, output.contrasts)
```
Arguments

\textbf{x1} \hspace{1em} \text{Factor variable (with contrasts based on weighted effect coding)}

\textbf{x2} \hspace{1em} \text{Factor variable (with contrasts based on weighted effect coding) or interval or ratio variable.}

\textbf{output.contrasts} \hspace{1em} \text{Specifies whether the contrast matrix of the interaction should be returned. Defaults to FALSE, returning the model matrix. Option currently only implemented for interactions between one weighted effect coded and one interval or ratio variable.}

Value

Returns a model matrix or contrast matrix for the interaction terms of (a.) two weighted effect coded variables, or (b.) one weighted effect coded and one interval or ratio variable.

Note

It should be noted that the procedure of applying weighted effect coding with interactions differs from the conventional way to apply contrasts in R. This is because the contrast matrix of the interaction differs from the multiplication of the contrast matrix/matrices of the interacted variables.

Author(s)

Rense Nieuwenhuis, Manfred te Grotenhuis, Ben Pelzer, Alexander Schmidt, Ruben König, Rob Eisinga

References


See Also

\textit{contr.wec}
Examples

data(BMI)

# Interaction two weighted effect coded categorical variables
BMI$childless.wec.yes <- BMI$childless.wec.no <- BMI$childless
contrasts(BMI$childless.wec.yes) <- contr.wec(BMI$childless, omitted="yes")
contrasts(BMI$childless.wec.no) <- contr.wec(BMI$childless, omitted="no")

BMI$age.wec.young <- BMI$age.wec.older <- BMI$age
contrasts(BMI$age.wec.young) <- contr.wec(BMI$age_categorical, omitted="Young (18-30)")
contrasts(BMI$age.wec.older) <- contr.wec(BMI$age_categorical, omitted="Older (60-70)")

model3a <- lm(BMI ~ childless.wec.yes + age.wec.young, data=BMI)
model3b <- lm(BMI ~ childless.wec.no + age.wec.older, data=BMI)
summary(model3a)
summary(model3b)

# Interaction
BMI$interact_c <- wec.interact(BMI$childless.wec.yes, BMI$age.wec.young)
BMI$interact_d <- wec.interact(BMI$childless.wec.yes, BMI$age.wec.older)
BMI$interact_e <- wec.interact(BMI$childless.wec.no, BMI$age.wec.young)
BMI$interact_f <- wec.interact(BMI$childless.wec.no, BMI$age.wec.older)

model3c <- lm(BMI ~ childless.wec.yes + age.wec.young + interact_c, data=BMI)
model3d <- lm(BMI ~ childless.wec.yes + age.wec.older + interact_d, data=BMI)
model3e <- lm(BMI ~ childless.wec.no + age.wec.young + interact_e, data=BMI)
model3f <- lm(BMI ~ childless.wec.no + age.wec.older + interact_f, data=BMI)
summary(model3c)
summary(model3d)
summary(model3e)
summary(model3f)

# Interaction weighted effect coded categorical variable and ratio/interval variable

data(PUMS)

PUMS$race.wec <- factor(PUMS$race)
contrasts(PUMS$race.wec) <- contr.wec(PUMS$race.wec, "White")
contrasts(PUMS$race.wec)

m.wec <- lm(wage ~ race.wec, data=PUMS)
summary(m.wec)

PUMS$race.educint <- wec.interact(PUMS$race.wec, PUMS$education.int)
m.wec.educ <- lm(wage ~ race.wec + education.int + race.educint, data=PUMS)
summary(m.wec.educ)
Index

*Topic **category**
- contr.wec, 3
- wec.interact, 5

*Topic **classes**
- contr.wec, 3
- wec.interact, 5

*Topic **datasets**
- BMI, 2
- PUMS, 4

*Topic **manip**
- contr.wec, 3
- wec.interact, 5

BMI, 2

contr.wec, 3, 6

PUMS, 4

wec.interact, 4, 5