Package ‘chest’

March 23, 2023

Type Package

Title Change-in-Estimate Approach to Assess Confounding Effects

Version 0.3.7

Description Applies the change-in-effect estimate method to assess confounding effects in medical and epidemiological research (Greenland & Pearce (2016) <doi:10.1146/annurev-publhealth-031914-122559> ). It starts with a crude model including only the outcome and exposure variables. At each of the subsequent steps, one variable which creates the largest change among the remaining variables is selected. This process is repeated until all variables have been entered into the model (Wang Z. Stata Journal 2007; 7, Number 2, pp. 183–196). Currently, the ‘chest’ package has functions for linear regression, logistic regression, negative binomial regression, Cox proportional hazards model and conditional logistic regression.

License GPL-2

Encoding UTF-8

LazyData true

Imports broom, ggplot2, survival, grid, forestplot, MASS, tibble, dplyr

Depends R (>= 2.20)

RoxygenNote 7.2.3

Suggests spelling, knitr, rmarkdown

VignetteBuilder knitr

Language en-US

NeedsCompilation no

Author Zhiqiang Wang [aut, cre]

Maintainer Zhiqiang Wang <menzies.uq@gmail.com>

Repository CRAN

Date/Publication 2023-03-23 09:50:13 UTC
\section*{Description}

In clinical trials and epidemiological studies, the association between an exposure and the outcome of interest in a study can be estimated by \textit{regression coefficients}, \textit{odds ratios} or \textit{hazard ratios} depending on the nature of study designs and outcome measurements. We use a general term \textit{effect estimate} here for any of those measurements in this document. Based on those measurements, we determine if a treatment is effective (or detrimental) or a factor is a risk factor. Imbalanced distributions of other factors could bias the effect estimates, called \textit{confounding}. One way to assess the confounding effect of a factor is to examine the difference in effect estimates between models with and without a specific factor. '\texttt{chest}' allows users quickly calculate the changes when potential confounding factors are sequentially added to the model in a stepwise fashion. At each step, one variable which creates the largest change (\%) of the effect estimate among the remaining variables is added to the model. '\texttt{chest}' returns a graph and a data frame (table) with effect estimates (95\% CI) and change (\%) values. The package currently has the following main functions: '\texttt{chest_lm}' for linear regression, '\texttt{chest_glm}' for logistic regression and Poisson regression, '\texttt{chest_clogit}' for matched logistic regression, '\texttt{chest_nb}' for negative binomial regression and '\texttt{chest_cox}' for Cox proportional hazards models.

\section*{References}


\section*{Examples}

\begin{verbatim}
? chest_glm
? chest_cox
? chest_clogit
? chest_lm
? chest_nb
? chest_plot
? chest_forest
\end{verbatim}
Assessing confounding effects using conditional logistic regression models

Description

'chest_clogit' is used to fit many Conditional Logistic Regression models to assess confounding effects.

Usage

chest_clogit(
  crude,
  xlist,
  data,
  method = "exact",
  na_omit = TRUE,
  plus = "+ ",
  indicate = FALSE,
  ...
)

Arguments

- **crude**: An object of formula for the initial model, generally crude model. However, any other variables can also be included here as the initial model.
- **xlist**: A vector of characters with all variable names of potential confounders.
- **data**: Data frame.
- **method**: See 'clogit', default is the "exact" method.
- **na_omit**: Remove all missing values, default: 'na_omit = TRUE'.
- **plus**: Change the + sign before variable names.
- **indicate**: indicate the calculation progress.
- **...**: Further optional arguments.

Value

A table with effect estimates and their changes at all steps.

See Also

- chest
- 'clogit' in 'survival'
Examples

vlist <- c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")
chest_clogit(
  crude = "Endpoint ~ Diabetes + strata(mid)",
  xlist = vlist, data = diab_df
)

### Description

'chest_cox' is used to assess confounding effects using Proportional Hazards Regression Model ('coxph' from 'survival' package). It presents the effect estimates (such as hazard ratios) for the association between exposure and outcome variables by adding other variables (potential confounders) to the model sequentially. The order of variables to be added is based on the magnitudes of the changes in effect estimates.

### Usage

chest_cox(
  crude,
  xlist,
  data,
  na_omit = TRUE,
  plus = " + ",
  indicate = FALSE,
  ...
)

### Arguments

**crude**
An object of *formula* for initial model, generally crude model. However, any other variables can also be included here as the initial model.

**xlist**
A *vector* of characters with variable names of potential confounders.

**data**
*Data frame*.

**na_omit**
Remove all missing values, default: 'na_omit = TRUE'.

**plus**
Change the + sign before variable names.

**indicate**
indicate the progress.

**...**
Further optional arguments for forestplot.

### Value

A table with effect estimates and their changes at all steps.
**chest_forest**

**Description**

`chest_forest` plots effect estimates and change-in-estimate values with forestplot package.

**Usage**

```r
chest_forest(
  data,
  var_lab = "Variables",
  est_lab = "Estimate (95% CI)",
  change_lab = "Change, %",
  digits = ".2f",
  digits_change = ".1f",
  hrzl_lines = gpar(col = "#444444"),
  plus = " + ",
  ...
)
```

**Arguments**

- **data** *Object* from `chest_cox`, `chest_glm`, `chest_lm`, `chest_clogit`, or `chest_nb`, including effect estimate values and change-in-estimate values.
- **var_lab** *Character* string for the column name of variables in the graph.
- **est_lab** *Character* string for the column name of effect estimates.
- **change_lab** *Character* string for the column name of "Changes".
- **digits** Set the display format for number in the graph other than the "Change" column. Default: ".2f"
- **digits_change** Set the format for the "Change" column. Default: ".1f"
- **hrzl_lines** A *logic* to include or remove horizontal line.
- **plus** Change the + sign before variable names.
- **...** Further optional arguments for forestplot.

**Examples**

```r
vlist <- c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")

chest_cox(crude = "Surv(t0, t1, Endpoint) ~ Diabetes", xlist = vlist, data = diab_df)
```

---

**See Also**

'survival'

---

**Description**

Plot effect estimates and change-in-estimate values (forestplot type)
Value

A table with effect estimates and their changes at all steps.

See Also

'forestplot'

Examples

vlist <- c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")
results <- chest_cox(crude = "Surv(t0, t1, Endpoint) ~ Diabetes", xlist = vlist, data = diab_df)
chest_forest(results)

chest_glm Assessing confounding effects using Generalized Linear regression models

Description

Please note: 'chest_glm' is used to assess confounding effects using Generalized Linear Models, such as logistic regression and Poisson regression with 'glm'. It presents odds ratios or rate ratios for the association between exposure and outcome variables by adding other variables (potential confounders) to the model sequentially. The order of variables to be added is based on the magnitudes of the changes in effect estimates.

Usage

chest_glm(
  crude,
  xlist,
  data,
  family = "binomial",
  method = "glm.fit",
  na_omit = TRUE,
  indicate = FALSE,
  plus = " + ",
  ...
)

Arguments

  crude  An object of formula for initial model, generally crude model. However, any other variables can also be included here as the initial model.
  xlist  A vector of characters with all variable names of potential confounders.
  data   Data frame.
  family Description of the error distribution. Default is "binomial".
Method to detect for singularity.

Remove all missing values, default: 'na_omit = TRUE'.

indicate progress.

Change the + sign before variable names.

Further optional arguments for forestplot.

Value

A table with effect estimates and their changes at all steps.

See Also

'glm' 'stats'

Examples

chest_glm(
  crude = "Endpoint ~ Diabetes", xlist = c("Age", "Sex", "Married"),
  na_omit = TRUE, data = diab_df
)

chest_lm

Assessing confounding effects using Linear regression models

Description

'chest_lm' is used to assess confounding effects using Linear Regression Models. It presents linear regression coefficients as effect estimates and changes when other variables are added sequentially to the model.

Usage

chest_lm(
  crude,
  xlist,
  data,
  method = "qr",
  na_omit = TRUE,
  indicate = FALSE,
  plus = " + ",
  ...
)
Arguments

- **crude**: An object of *formula* for initial model, generally crude model. However, any other variables can also be included here as the initial model.
- **xlist**: A *vector* of characters with all variable names of potential confounders.
- **data**: *Data frame*.
- **method**: The method to be used; see *'lm'*.
- **na_omit**: Remove all missing values.
- **indicate**: Indicate progress
- **plus**: Change the + sign before variable names.
- **...**: Further optional arguments.

Value

A table with effect estimates and their changes at all steps.

See Also

- *'lm'* of *'stats'*

Examples

```r
vlist <- c("Age", "Sex", "Married", "Cancer", "CVD", "Education", "Income")
chest_lm(crude = "BMI ~ Diabetes", xlist = vlist, data = diab_df, na_omit = TRUE)
```

---

**chest_nb**

Assessing confounding effects using Negative Binomial regression models

Description

It fits negative binomial regression models to present *rate ratios* for the association between exposure and outcome variables by adding other variables (potential confounders) to the model sequentially. The order of variables to be added is based on the magnitudes of the changes in effect estimates.

Usage

```r
chest_nb(
  crude, 
  xlist, 
  data, 
  method = "glm.fit",
  na_omit = TRUE,
  indicate = FALSE,
  plus = " + ",
  ...
)
```
Arguments

- **crude**: An object of *formula* for initial model, generally crude model. However, any other variables can also be included here as the initial model.
- **xlist**: A *vector* of characters with all variable names of potential confounders.
- **data**: *Data frame*.
- **method**: Method to detect for singularity.
- **na_omit**: Remove all missing values, default: `na_omit = TRUE`.
- **indicate**: indicate progress.
- **plus**: Change the + sign before variable names.
- **...**: Further optional arguments.

Value

A table with effect estimates and their changes at all steps.

See Also

`glm.nb` of `MASS`

Examples

```r
library(MASS)
df <- quine
results <- chest_nb("Days ~ Lrn", xlist = c("Sex", "Age", "Eth"), data = df)
results$data
```

---

 chest_plot  Plot effect estimate and change-in-estimate values (ggplot type)

Description

Plot effect estimate and change-in-estimate values (ggplot type)

Usage

```r
chest_plot(
  data,
  no_values = FALSE,
  ylab = NULL,
  xlab = NULL,
  change_lab = "Change, \%",
  digits = "%.2f",
  digits_change = "%.1f",
  plus = " + ",
  nudge_y = 0.4,
```
Arguments

data: Object from `chest_cox`, `chest_glm`, `chest_lm`, `chest_clogit`, or `chest_nb`, including effect estimate values and change-in-estimate values.

no_values: Suppress effect estimate values in plot, default is FALSE.

ylab: Add y label.

xlab: Add x label.

change_lab: Character string for the column name of "Change" in the graph

digits: Set the display format for number in the graph other than the "Change" column. Default: "%.2f"

digits_change: Set the format for the "Change" column. Default: "%.1f"

plus: Change the + sign before variable names.

nudge_y: Adjust vertical distance between values and point marker.

nudge_x: Adjust horizontal distance between values and point marker.

hjust: Adjust horizontal alignment.

height: Change the height of error bars.

point_size: Change point marker size.

point_shape: Change point marker shape.

dline_type: Change vertical line type.

dline_color: Change vertical line color.

vbar_color: Change error bar color.

zero: x-axis coordinate for vertical non-effect line, see `forestplot`.

value_position: Change the position of value labels.

... Further optional arguments for `forestplot`.

Value

a `ggplot` object.
See Also

'ggplot2'

Examples

vlist <- c("Age", "Sex", "Married", "Education", "Income")
results <- chest_glm(
  crude = "Endpoint ~ Diabetes",
  xlist = vlist, na_omit = TRUE, data = diab_df
)
chest_plot(results)

---

**diab_df**

Example data: Health outcomes of 2372 adults with and without diabetes

**Description**

A data frame with 2372 rows and 14 variables with diabetes status diabetes and mortality status endpoint. For the purpose of demonstrate, assume that we are interested in the association between diabetes and endpoint. Other variables are considered as possible confounders. The purposes of this dataset is to illustrate those functions in chest package only. Therefore, we assume it is a cohort design for Cox Proportional Hazard regression, and a case-control design for logistic regression.

**Usage**

diab_df

**Format**

A data frame with 2372 rows and 14 variables:

- **Diabetes** diabetes status 1: with diabetes 0: without diabetes
- **Endpoint** mortality status 1: reached end point, and 0: survived
- **Age** Age, in years
- **Sex** sex, 1: male, 2: Female
- **BMI** Body mass index
- **Married** marital status 1: married, 0: not
- **Smoke** smoking status 1: smoker, 0: non-smoker
- **CVD** cardiovascular disease 1: yes 0: no
- **Cancer** cancer 1: yes, 0: no
- **Education** education 1: high, 0: low
- **Income** income 1: high, 0: low
- **t0** time (age) at the start of the follow-up
- **t1** time (age) at the end of the follow-up
- **mid** matched set id, for conditional logistic regression
Index

* datasets
  diab_df, 11
  chest, 2
  chest_clogit, 3
  chest_cox, 4
  chest_forest, 5
  chest_glm, 6
  chest_lm, 7
  chest_nb, 8
  chest_plot, 9
  diab_df, 11