Package 'ezfun'

August 6, 2018

Title Emily C. Zabor's functions

Version 0.1.3

Date 2016-12-10

Description This package contains a number of functions that generate and format results of common procedures for clincial projects into table form for printing in R Markdown Word documents. A few basic utility functions for common procedures are also included.

Depends R (>= 3.1.0)

License GPL-2

LazyData TRUE

Imports survival, aod, cmprsk, clinfun, shiny, lme4

RoxygenNote 6.0.1

Author Emily Zabor [aut, cre]

Maintainer Emily Zabor <zabore@mskcc.org>

R topics documented:

`bycont` *Table of one or more categorical variables by a single continuous variable*

Description

`bycont` takes a list of categorical variabls and returns median(min, max) of the single continuous variable within each level of each categorical variable. Computes Kruskall-Wallis p-values.

Usage

```
bycont(catvars, contvar, dat, pval = TRUE)
```

Arguments

`catvars`	is a list of the categorical variables for the rows of the table e.g. list('Gene1', 'Gene2')
`contvar`	is the continuous variable you want summarized by each categorical variable. Must be in quotes.
`dat`	is the dataset to use for analysis
`pval`	takes the value TRUE or FALSE indicating whether p-values should be computed. Defaults to TRUE. When TRUE, Kruskal-Wallis p-values are produced.

Value

Returns a dataframe

Author(s)

Emily C Zabor <zabore@mskcc.org>

`ez_pal` *A custom color scale made by Emily Zabor, with help from Coolors app*

Description

Basically copying code from hrbrmstr/hrbrthemes colors.r

Usage

```
ez_pal()
```

Examples

```
library(scales)
scales::show_col(ez_pal()(9))
```

lowerchar *Convert to lowercase*

Description

lowerchar converts the levels of character variables from upper or mixed case to lower case

Usage

```
lowerchar(dfname)
```

Arguments

dfname	is the name of the dataframe on which to perform the action

Value

Nothing is returned from lowerchar, the action is simply perfomed on the columns of dataframe dfname

Author(s)

Emily C Zabor <zabore@mskcc.org>

msk_palette *MSK branded colors palette generator*

Description

MSK branded colors palette generator

Usage

```
msk_palette(name, n, type = c("discrete", "continuous"))
```

Arguments

name	Name of desired palette. Choices are: all, main (default), blues, oranges, contrast
n	Number of colors desired. If omitted, uses all colours.
type	Either "continuous" or "discrete". Use continuous if you want to automatically interpolate between colours. @importFrom graphics rgb rect par image text

Value

A vector of colours.

Examples

```
msk_palette("main")
msk_palette("blues")
msk_palette("contrast")

library(ggplot2)

# use a single brand color from a palette
ggplot(mtcars, aes(hp, mpg)) +
geom_point(size = 4, color = msk_palette("main")[1])

# use a discrete color scale
ggplot(iris, aes(Sepal.Width, Sepal.Length, color = Species)) +
geom_point(size = 4) +
scale_color_manual(values = msk_palette("contrast"))

# use a continuous color scale
ggplot(iris, aes(Sepal.Width, Sepal.Length, color = Sepal.Length)) +
geom_point(size = 4, alpha = .6) +
scale_color_gradientn(colors = msk_palette("blues", type = "continuous"))
```

---

msk_palettes *Complete list of MSK branded color palettes*

Description

Complete list of MSK branded color palettes

Usage

```
msk_palettes
```

Format

An object of class `list` of length 5.

---

mvcoxres *Format resuls from multivariable Cox regression model*

Description

`mvcoxres` takes a multivariable Cox regression object and formats the resulting HR (95% CI) and p-values into a table

Usage

```
mvcoxres(mod)
```

Arguments

`mod`	is a multivariable Cox regression object

Value

Returns a dataframe

Author(s)

Emily C Zabor <zabore@mskcc.org>

`mvcrrres` *Format resuls from multivariable competing risks regression model*

Description

`mvcrrres` takes a multivariable competing risks regression object and puts the resulting HR (95% CI) and p-values into a table

Usage

```
mvcrrres(mod)
```

Arguments

`mod`	is a multivariable Cox regression object

Value

Returns a dataframe

Author(s)

Emily C Zabor <zabore@mskcc.org>

`mvlmres` *Format resuls from multivariable linear regression model*

Description

`mvlmres` takes a multivariable linear regression object and formats the resulting Est (SE) and p-values into a table

Usage

```
mvlmres(mod)
```

Arguments

`mod`	is a multivariable linear regression object from `lm`

Value

Returns a dataframe

Author(s)

Emily C Zabor <zabore@mskcc.org>

---

mvlogitres *Format resuls from multivariable logistic regression model*

Description

`mvlogitres` takes a multivariable logistic regression object and formats the resulting OR (95% CI) and p-values into a table

Usage

```
mvlogitres(mod)
```

Arguments

`mod`	is a multivariable logistic regression object from `glm`

Value

Returns a dataframe

Author(s)

Emily C Zabor <zabore@mskcc.org>

---

ph2simonApp *Interactive Simon's 2-stage Shiny app*

Description

`ph2simonApp` is simply a Shiny interface for the `ph2simon` function from the `clinfun` package. No arguments need to be passed to the function.

Usage

```
ph2simonApp()
```

Value

The output includes 1) fields to enter the design parameters for the Simon 2-stage Phase II design, 2) R output with `ph2simon` results, 3) a paragraph interpreting the results, and 4) a plot of maximum versus expected number of patients indicating the optimal and minimax results.

scale_colour_ez *Discrete color & fill scales based on the ez palette*

Description

See ez_pal.

Usage

```
scale_colour_ez(...)

scale_color_ez(...)

scale_fill_ez(...)
```

Arguments

...	Other arguments passed on to discrete_scale to control name, limits, breaks, labels and so forth.

sdp *Get p-value from survdiff()*

Description

sdp returns the p-value from the survdiff function

Usage

```
sdp(sd)
```

Arguments

sd	is a survdiff object

Value

Returns a p-value rounded to 3 digits or "<.001" if the p-value is <.001

Author(s)

Emily C Zabor <zabore@mskcc.org>

tab1 *Table 1*

Description

tab1 takes lists of continuous and/or categorical variables and returns Median (spread) for continuous variables and N (%) for categorical variables. Produces a table with both an overall column and columns by another variable.

Usage

```
tab1(contvars, catvars, byvar, dat, col = TRUE, spread = "range",
  pval = TRUE, fisher = TRUE)
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age'). Can be NULL.
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race'). Can be NULL.
byvar	is the categorical variable you want to tabulate by across the columns (needs to be in quotes). Can be NULL.
dat	is the dataset to use for analysis
col	takes the value TRUE or FALSE indicating whether you want column percent (TRUE, default) or row percent (FALSE)
spread	takes the value "range" or "iqr" indicating whether you want (min, max) or (Q1, Q3) in summaries of continuous variables. Defaults to "range".
pval	takes the value TRUE or FALSE indicating whether p-values should be included. Defaults to TRUE. If TRUE, kruskal.test p-values are produced for continuous variables and either fisher.test or chisq.test p-values are produced for categorical variables. See param for testing details for categorical variables.
fisher	takes the value TRUE or FALSE. If TRUE, fisher.test p-values are produced. If FALSE, chisq.test p-values are produced.

Value

Returns a dataframe. If there are warnings or errors from kruskal.test, fisher.test, or chisq.test then NA is returned in place of the p-value.

Author(s)

Emily C Zabor <zabore@mskcc.org>

tab1_re *Table 1 with random effects model p-values*

Description

tab1_re takes lists of continuous and/or categorical variables and returns Median (spread) for continuous variables and N (%) for categorical variables. Produces a table with both an overall column and columns by another variable. For a binary by variable only, it produces p-values from a random effects model.

Usage

```
tab1_re(contvars, catvars, byvar, re, dat, col = TRUE, spread = "range")
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age'). Can be NULL.
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race'). Can be NULL.
byvar	is the categorical variable you want to tabulate by across the columns (needs to be in quotes). MUST BE 0/1 since it will be used as the outcome variable in glmer.
dat	is the dataset to use for analysis
col	takes the value TRUE or FALSE indicating whether you want column percent (TRUE, default) or row percent (FALSE)
spread	takes the value "range" or "iqr" indicating whether you want (min, max) or (Q1, Q3) in summaries of continuous variables. Defaults to "range".

Value

Returns a dataframe. If there are warnings or errors from glmer then NA is returned in place of the p-value.

Author(s)

Emily C Zabor <zabore@mskcc.org>

tabna *Cross-tabulation with* useNA = "ifany"

Description

tabna is an implementation of table with argument useNA = "ifany"

Usage

```
tabna(...)
```

Arguments

...	the function takes any standard arguments to table

Details

See the help file for table for detailed information about possible arguments to the function

Value

tabna returns a contingency table with NAs included, if any

Author(s)

Emily C Zabor <zabore@mskcc.org>

---

theme_ezbasic *Basic theme preferences for ggplot*

Description

Basic theme preferences for ggplot. Functionally a crib of https://github.com/hrbrmstr/hrbrthemes/blob/master/R/theme-ipsum.r with some default changes

Usage

```
theme_ezbasic(base_size = 11.5, plot_title_size = 16,
  plot_title_face = "plain", plot_title_margin = 10, subtitle_size = 13,
  subtitle_face = "plain", subtitle_margin = 15, strip_text_size = 12,
  strip_text_face = "plain", caption_size = 9, caption_face = "plain",
  caption_margin = 10, axis_text_size = base_size, axis_title_size = 9,
  axis_title_face = "plain", axis_title_just = "lt",
  plot_margin = margin(10, 10, 10, 10), grid_col = "#cccccc", grid = TRUE,
  axis_col = "#cccccc", axis = FALSE, ticks = FALSE,
  legend_title = FALSE, legend_bottom = TRUE, legend_just = "left")
```

Arguments

base_size	base font size
plot_title_face, plot_title_size, plot_title_margin	plot title face, size and margin
subtitle_face, subtitle_size	plot subtitle face and size
subtitle_margin	plot subtitle margin bottom (single numeric value)
strip_text_face, strip_text_size	facet label font face and size
caption_face, caption_size, caption_margin	plot caption face, size and margin
axis_text_size	font size of axis text

axis_title_face, axis_title_size	axis title font face and size
axis_title_just	axis title font justification, one of '[blmcrt]'
plot_margin	plot margin (specify with [ggplot2::margin()])
grid_col, axis_col	grid & axis colors; both default to '#cccccc'
grid	panel grid ('TRUE', 'FALSE', or a combination of 'X', 'x', 'Y', 'y')
axis	add x or y axes? 'TRUE', 'FALSE', "'xy'"
ticks	ticks if 'TRUE' add ticks
legend_title	includes legend title if 'TRUE', defaults to 'FALSE'
legend_bottom	places legend at bottom if 'TRUE', places legend to default position if 'FALSE'
legend_just	legend justification, one of 'right', 'left', 'center'. Defaults to 'left'.

`todate` *Convert to date format*

Description

`todate` converts any POSIXct format variables in the dataframe to date format

Usage

```
todate(dfname)
```

Arguments

dfname	is the name of the dataframe on which to perform the action

Details

Note that this function will mainly apply to dataframes imported using the `read_excel` function from the `readxl` package. Dataframes imported using, for example, `read.csv` instead will have dates in character format and therefore `todate` will not apply.

Value

Nothing is returned from `todate`, the action is simply perfomed on the columns of dataframe `dfname`

Author(s)

Emily C Zabor <zabore@mskcc.org>

uvcoxph *Table of univariable Cox regression results*

Description

uvcoxph takes lists of continuous and/or categorical variables, runs a univariable coxph model for each, and puts the resulting HR (95% CI) and p-value into a table suitable for printing in a Word R Markdown file.

Usage

```
uvcoxph(contvars, catvars, event, time, dat, strata = NULL)
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age')
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race')
event	is the event indicator (needs to be in quotes)
time	is the survival time variables (needs to be in quotes)
dat	is the dataset for analysis
strata	is a possible strata term for use in calculating the log-rank p-values. Defaults to NULL. Entries should be in quotes, e.g. "Surgeon"

Value

Returns a dataframe. If there are warnings or errors from coxph then blank rows are returned.

Author(s)

Emily C Zabor <zabore@mskcc.org>

uvcrr *Table of univariable competing risks regression results*

Description

uvcrr takes lists of continuous and/or categorical variables, runs a univariable crr model for each, and puts the resulting HR (95% CI) and p-value into a table suitable for printing in a Word R Markdown file.

Usage

```
uvcrr(contvars, catvars, event, time, dat)
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age')
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race')
event	is the event indicator (needs to be in quotes)
time	is the survival time variables (needs to be in quotes)
dat	is the dataset for analysis

Details

uvcrr uses all function defaults to crr. For example, the failure code is set to 1. See the help file for crr for additional details.

Value

Returns a dataframe. If there are warnings or errors from crr then blank rows are returned.

Author(s)

Emily C Zabor <zabore@mskcc.org>

---

uvlm *Table of univariable linear regression results*

Description

uvlm takes lists of continuous and/or categorical variables, calls lm to run a linear regression model for each, and returns a table with Est (SE) and p-value for each variable that is suitable for printing in a Word R Markdown file.

Usage

```
uvlm(contvars, catvars, out, dat)
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age')
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race')
out	is the continuous outcome variable (needs to be in quotes)
dat	is the dataset for analysis

Value

Returns a dataframe. If there are warnings or errors from lm then blank rows are returned.

Author(s)

Emily C Zabor <zabore@mskcc.org>

uvlogit *Table of univariable logistic regression results*

Description

uvlogit takes lists of continuous and/or categorical variables, calls glm to run a logistic regression model for each, and returns a table with OR (95 Word R Markdown file.

Usage

```
uvlogit(contvars, catvars, out, dat)
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age')
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race')
out	is the binary outcome variable. Must be coded 0/1. (needs to be in quotes)
dat	is the dataset for analysis

Value

Returns a dataframe. If there are warnings or errors from glm then blank rows are returned.

Author(s)

Emily C Zabor <zabore@mskcc.org>

uvsurv *Table of univariable survival analysis results*

Description

uvsurv takes lists of continuous and/or categorical variables. For continuous variables, coxph returns HR (95% CI) and log-rank p-values. For categorical variables, coxph returns HR (95% CI) and log-rank p-values and survfit produces median survival (95% CI) and a survival estimate at a specified time. Results are put into a table suitable for printing in a Word R Markdown file.

Usage

```
uvsurv(contvars, catvars, event, time, test, dat, strata = NULL)
```

Arguments

contvars	is a list of the continuous variables you want in the rows e.g. list('Age')
catvars	is a list of the categorical variables you want in the rows e.g. list('Gender','Race')
event	is the survival event indicator (needs to be in quotes)
time	is the survival time variable (needs to be in quotes)
test	is the timepoint you would like to estimate, in whatever units the survival time is in
dat	is the dataset to use for analysis
strata	is a possible strata term for use in calculating the log-rank p-values. Defaults to NULL. Entries should be in quotes, e.g. "Surgeon"

Value

Returns a dataframe

Author(s)

Emily C Zabor <zabore@mskcc.org>

Index