Package 'Table1'

May 16, 2016

Type Package

Title Summary Table
Version 1.0
Date 2015-06-18
Author Erica Wozniak
Maintainer Erica Wozniak <ewozniak@mcw.edu></ewozniak@mcw.edu>
Description This package creates data frames of formatted summaries for categorical and continuous variables. Basic significance tests across strata are available, in addition to several one-sample tests. Options to output results to LaTeX, HTML, and plain text to the console are currently supported.
Imports xtable, htmlTable
Suggests survival
License GPL-2
NeedsCompilation no
R topics documented:
Table1-package
cat.var
cont.var
make.table
out.html
out.latex
out.plain
quick.table
stat.col
Index 17

2 Table1-package

|--|

Description

This package creates data frames of formatted summaries for categorical and continuous variables. Basic significance tests across strata are available, in addition to several one-sample tests. Options to output results to LaTeX, HTML, and plain text to the console are currently supported.

Details

Package: Table1
Type: Package

Title: Summary Table

Version: 1.0

Date: 2015-06-18 Author: Erica Wozniak

Maintainer: Erica Wozniak <ewozniak@mcw.edu>

Description: This package creates data frames of formatted summaries for categorical and continuous variables. Basic s

Imports: xtable, htmlTable

Suggests: survival License: GPL-2

The main summary functions in this package are make.table and quick.table. The make.table function accepts separate vectorized inputs for categorical and continuous variables and provides numerous options to customize table output, including significance testing. By contrast, quick.table allows for very little customization; it is instead intended to generate rapid summaries of datasets with minimal input, dynamically classifying each variable as either categorical or continuous.

Index of help topics:

Table1-package Summary Table

cat.var Categorical Variable Summary cont.var Continuous Variable Summary

make.table Make Summary Table

out.html HTML Output
out.latex LaTeX Output
out.plain Plain Text Output
quick.table Quick Summary Table
stat.col Statistical Testing

Note

You must clean (e.g., replace blank entries with NAs) the data frame of interest before applying these functions.

Working installations of xtable and htmlTable are required for LaTeX and HTML output, respectively.

cat.var 3

Author(s)

Erica Wozniak

Maintainer: Erica Wozniak <ewozniak@mcw.edu>

See Also

make.table quick.table cat.var cont.var stat.col out.plain out.latex out.html

Examples

```
library(survival)
library(htmlTable)
attach(pbc)
#Plain text table generated with make.table using vectorized inputs
make.table(dat = pbc,
   cat.varlist = c('stage', 'sex'),
    cat.header = c('Stage', 'Sex'),
   cat.rownames = list(c('I', 'II', 'III', 'IV'), c('Male', 'Female')),
   cat.ptype = c('fisher', 'fisher'),
cont.varlist = c('bili', 'copper'),
    cont.header = c('Bilirubin', 'Urine copper'),
    cont.ptype = c('wilcox', 'wilcox'),
           = c('trt'),
    # Remove row percents from all categorical variables (list is recycled)
    cat.rmstat = list('row'),
    # Remove missing, min/max from bili and no summary stats from copper
    cont.rmstat = list(c('miss', 'minmax'), 'None'),
    colnames = c('', 'D-penicillamine', 'Placebo', 'Overall', 'p-value')
    )
# HTML table generated with quick.table
library(htmlTable)
quick.table(dat = pbc,
   strat = 'sex',
colnames = c('', 'Male', 'Female', 'Overall'),
    # See www.color-hex.com for codes
    output = 'html'
    stripe.col = '#d0f2b0'
```

cat.var

Categorical Variable Summary

Description

This function produces counts, row percentages, column percentages, and counts of missing values for each level of a categorical variable input, optionally stratified by another variable, and with the option to compute p-values.

Usage

```
cat.var(var, strat, dec, rownames, header, ptype, pname, cat.rmstat, vspace, output)
```

4 cat.var

Arguments

Categorical variable input. Required. var Column stratifying variable. Optional. strat dec Number of decimal places on summary statistics. Optional; defaults to 2. Ordered character vector of row names for each level of 'var'. Optional; defaults rownames to the value of each factor level. Character vector providing a short variable description. Optional; defaults to header variable name. Statistical test for p-value. Optional; defaults to none. See stat.col for a ptype complete list of options. If TRUE then the name of the statistical test used will be printed below the pname p-value. Optional; defaults to FALSE. cat.rmstat Vector naming the statistics that should be removed from the summary data frame. Optional; defaults to none. "count" Remove overall count/percent "row" Remove row percents, leaving only column percents "col" Remove column percents, leaving only row percents

TRUE for additional vertical whitespace between variables (default).

Value

vspace output

Returns a data frame of categorical summary statistics.

Author(s)

Erica Wozniak

See Also

Table1 cont.var make.table quick.table stat.col out.plain out.latex out.html

"miss" Remove summary of missing values

Output type passed from make. table.

Examples

cont.var 5

cont.var

Continuous Variable Summary

Description

This function computes the mean (SD), median (IQR), Q1/Q3, min/max, and a count of missings for a continuous variable input, optionally stratified by another variable, and with the option to compute p-values.

Usage

```
cont.var(var, strat, dec, header, ptype, pname, cont.rmstat, vspace, output)
```

Arguments

output

var	Continuous variable input. Required.
strat	Column stratifying variable. Optional.
dec	Number of decimal places on summary statistics. Optional; defaults to 2.
header	Character vector providing a short variable description. Optional; defaults to variable name.
ptype	Statistical test for p-value. Optional; defaults to none. See stat.col for a complete list of options.
pname	If TRUE then the name of the statistical test used will be printed below the p-value (default). Optional.
cont.rmstat	Vector naming the statistics that should be removed from the summary data frame. Optional; defaults to none.
	"count" Remove overall count
	"meansd" Remove mean and standard deviation
	"mediqr" Remove median and interquartile range
	"q1q3" Remove 25th and 75th percentiles
	"minmax" Remove minimum and maximum
	"miss" Remove summary of missing values
vspace	TRUE for additional vertical whitespace between variables (default). Optional.

Output type passed from make.table.

6 make.table

Details

Note that warnings and errors beyond those automatically provided by some base R functions are not in place for the improper use of significance testing. For example, if a t-test is requested across more than two strata, the code will execute and the returned p-value will reflect a test of only the first two strata.

Value

Returns a data frame of continuous summary statistics.

Author(s)

Erica Wozniak

See Also

Table1 cat.var make.table quick.table stat.col out.plain out.latex out.html

Examples

```
library(survival)
# Default output
cont.var(var = pbc$age,
        strat = pbc$sex
# Naming options and p-value
cont.var(var = pbc$age,
        strat = pbc$sex,
        header = "Age",
        ptype = "ttest"
         )
# Remove min/max and missing statistics, and test for p-value
cont.var(var
                    = pbc$age,
         strat
                    = pbc$sex,
                    = "Age",
        header
                = "ttest",
        ptype
        cont.rmstat = c("minmax", "miss"),
         pname
                    = FALSE
         )
```

make.table

Make Summary Table

Description

This function produces summary statistics for categorical and continuous variables and combines the results into a data frame for plain text, HTML, or LaTeX output, serving as a wrapper to apply vectorized inputs to cat.var and cont.var. The input dataset must be cleaned (e.g. blank cells and NaNs replaced with NAs).

make.table 7

Usage

make.table(dat, cat.varlist, cat.header, cat.rownames, cat.ptype,
cont.varlist, cont.header, cont.ptype, strat, cat.rmstat, cont.rmstat,
dec, pname, colnames, output, vspace, varorder, stripe, stripe.col, header.style,
factor.style, stat.style, nowrap, caption, footer, tspanner, n.tspanner,
cgroup, n.cgroup, col.columns)

Arguments

dat	Clean data frame input of all variables used for the summary table.
cat.varlist	Vector of categorical variable names. Optional; defaults to NULL.
cat.header	Ordered vector of short descriptions for each categorical variable. Optional; defaults to variable names.
cat.rownames	Ordered list of vectors providing row names for each level of each vector found in cat.varlist. Optional; defaults to the coded value of each level.
cat.ptype	Ordered vector of tests requested to add p-values to table for categorical variables. Optional; defaults to no p-values. See <pre>stat.col</pre> for a complete list of options.
cont.varlist	Vector of continuous variable names. Optional; defaults to NULL.
cont.header	Ordered vector of short descriptions for each continuous variable. Optional; defaults to variable name.
cont.ptype	Ordered vector of tests requested to add p-values to table for continuous variables. Optional; defaults to no p-values. See <pre>stat.col</pre> for a complete list of options.
strat	Vector of one or more stratifying variables. Optional; defaults to no stratification.
cat.rmstat	Ordered list of vectors of statistics that should be removed from respective final summaries of categorical variables.
cont.rmstat	Ordered list of vectors of statistics that should be removed from respective final summaries of continuous variables.
dec	Integer value indicating the number of decimal places to be printed for each summary statistic. Optional; defaults to 2.
pname	Logical indicator for whether to print statistical test used beneath each p-value. Optional; defaults to TRUE.
colnames	Ordered vector of column names for the table. Optional; defaults to levels of stratifying variable(s) with an overall summary column.
output	Choice of formatted table output: "html", "latex", or "plain". Optional; defaults to plain text printed to the console.
vspace	TRUE for additional vertical whitespace between variables (default). Optional.
varorder	Order of variables in the outputted table. Defaults to "data", or the order of variables in the inputted dataset. Alternatively, variables may be presented alphabetically using the "abc" option. Optional.
stripe	HTML output only: logical indicator for whether to zebra stripe every other variable in the table. Optional; defaults to TRUE.
stripe.col	HTML output only: hex color code to be used for zebra striping the table. Optional; defaults to a light gray, #F7F7F7.

8 make.table

header.style	Variable name style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "bold".
factor.style	$\label{thm:continuous} \begin{tabular}{ll} Variable factor levels style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "bold". \end{tabular}$
stat.style	Summary statistic style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "plain".
nowrap	HTML output only: option to prevent text wrapping of long lines to the table output. Useful for generating wide tables with knitr. Optional; defaults to TRUE.
caption	Table caption. Optional; no default.
footer	Table footer. Optional; no default.
tspanner	$HTML\ output\ only:\ Rowwise\ table\ spanner,\ passed\ to\ html Table.\ Optional;\ see\ ?html Table\ for\ default\ information.$
n.tspanner	HTML output only: Vector of rows in each table spanner, passed to htmlTable. Optional; see ?htmlTable for default information.
cgroup	$HTML\ output\ only:\ Column\ group\ headers,\ passed\ to\ html Table.\ Optional;\ see\ ?html Table\ for\ default\ information.$
n.cgroup	HTML output only: Vector of columns under each group header, passed to htmlTable. Optional; see ?htmlTable for default information.
col.columns	HTML output only: Vector of column colors, passed to htmlTable. Optional; defaults to "none".

Value

Returns a table of summary statistics for specified categorical and continuous variables in the requested output format. The plain text default returns a formatted data frame to the console.

Note

A few miscellaneous notes:

Table output to HTML and LaTeX require the package dependencies htmlTable and xtable

If missing values occur in stratifying variables (strat =), then the observations with missings will be removed from the table, and a comment will be printed indicating the total number removed. If multiple stratifying variables are provided, look to the console for a summary of total missing observations for each variable.

Author(s)

Erica Wozniak

See Also

Table1 quick.table cat.var cont.var stat.col out.plain out.latex out.html

out.html 9

Examples

```
library(survival)
# Example with non-HTML options
make.table(dat = pbc,
    cat.varlist = c('stage', 'sex'),
    cat.header = c('Stage', 'Sex'),
cat.rownames = list(c('I', 'II', 'III', 'IV'), c('Male', 'Female')),
cat.ptype = c('fisher', 'fisher'),
cont.varlist = c('bili', 'copper'),
    cont.header = c('Bilirubin', 'Urine copper'),
    cont.ptype = c('wilcox', 'wilcox'),
                 = c('trt'),
    # Remove row percents from all categorical variables (list is recycled)
    cat.rmstat = list('row'),
    \# Remove missing, min/max from bili and no summary stats from copper
    cont.rmstat = list(c('miss', 'minmax'), 'None'),
                 = c('', 'D-penicillamine', 'Placebo', 'Overall', 'p-value')
# Example with HTML options
library(htmlTable)
make.table(dat = pbc,
    cat.varlist = c('stage', 'sex'),
    cat.header = c('Stage', 'Sex'),
    cat.rownames = list(c('I', 'II', 'III', 'IV'), c('Male', 'Female')),
    cat.ptype = c('fisher', 'fisher'),
    cont.varlist = c('bili', 'copper'),
    cont.header = c('Bilirubin', 'Urine copper'),
    cont.ptype = c('wilcox', 'wilcox'),
    strat
                 = c('trt'),
    cat.rmstat = list('row'),
    colnames = c('',
output = 'html'
                 = c('', 'D-penicillamine', 'Placebo', 'Overall', 'p-value'),
    # Check out \url{www.color-hex.com}
    stripe.col = '#f4dfd0'
    )
```

out.html

HTML Output

Description

This function utilizes htmlTable to output the results of the make.table and quick.table functions to the R Viewer as a formatted HTML table, and may be used within a knitr document.

Usage

```
out.html(tab, colnames, stripe, stripe.col, header.style,
factor.style, stat.style, caption, footer, tspanner, n.tspanner,
cgroup, n.cgroup, col.columns, nowrap, vspace)
```

out.html

Arguments

tab	A make.table or quick.table generated data frame. Required.
colnames	A vector of column names for the table output. Optional.
stripe	Logical indicator for whether to include lightly shaded zebra striping of every other variable in the output. Optional; defaults to TRUE.
stripe.col	Hex code for the color of the zebra striping. Optional; defaults to a light gray, #F7F7F7.
header.style	Variable name style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "bold".
factor.style	Variable factor levels style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "bold".
stat.style	Summary statistic style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "plain".
caption	Table caption. Optional; no default.
footer	Table footer. Optional; no default.
tspanner	Rowwise table spanner, passed to htmlTable. Optional; see ?htmlTable for default information.
n.tspanner	Vector of rows in each table spanner, passed to htmlTable. Optional; see ?htmlTable for default information.
cgroup	Column group headers, passed to htmlTable. Optional; see ?htmlTable for default information.
n.cgroup	Vector of columns under each group header, passed to htmlTable. Optional; see ?htmlTable for default information.
col.columns	Vector of column colors, passed to htmlTable. Optional; defaults to "none".
nowrap	Option to prevent text wrapping of long lines to the table output. Useful for generating wide tables with knitr. Optional; defaults to TRUE.
vspace	TRUE for additional vertical whitespace between variables (default).

Details

A call to this function is an optional argument to both make.table and quick.table via output='html'.

Value

A formatted HTML table in the Viewer or code for inclusion in a knitr document.

Author(s)

Erica Wozniak

See Also

Table1 make.table quick.table cat.var cont.var stat.col out.plain out.latex

out.latex 11

Examples

```
library(survival)
library(htmlTable)

make.table(dat = pbc,
    cat.varlist = c('stage', 'sex'),
    cat.header = c('Stage', 'Sex'),
    cat.rownames = list(c('I', 'II', 'III', 'IV'), c('Male', 'Female')),
    cat.ptype = c('fisher', 'fisher'),
    cont.varlist = c('bili', 'copper'),
    cont.header = c('Bilirubin', 'Urine copper'),
    cont.ptype = c('wilcox', 'wilcox'),
    strat = c('trt'),
    cat.rmstat = list('row'),
    colnames = c('', 'D-penicillamine', 'Placebo', 'Overall', 'p-value'),
    output = 'html',
    # Check out \url{www.color-hex.com}
    stripe.col = '#f4dfd0'
    )
```

out.latex

LaTeX Output

Description

This function utilizes xtable to output summary table results to the console as formatted LaTeX code, and may be used within a Sweave or knitr document.

Usage

```
out.latex(tab, colnames, header.style, factor.style, stat.style)
```

Arguments

tab	A single call to either cat.var or cont.var, or a data frame produced via rbind of multiple calls to these functions. Required.
colnames	A vector of column names for the table output. Optional.
header.style	Variable name style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "bold".
factor.style	Variable factor levels style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "bold".
stat.style	Summary statistic style: "plain", "bold", "italic", or "bolditalic". Optional; defaults to "plain".

Details

A call to this function is an optional argument to make.table and quick.table via output='latex'.

Value

LaTeX code for a formatted table.

12 out.plain

Author(s)

Erica Wozniak

See Also

Table1 make.table quick.table cat.var cont.var stat.col out.plain out.html

Examples

```
library(xtable)
library(survival)
make.table(dat = pbc,
   cat.varlist = c('stage', 'sex'),
   cat.header = c('Stage', 'Sex'),
   cont.header = c('Bilirubin', 'Urine copper'),
   cont.ptype = c('wilcox', 'wilcox'),
         = c('trt'),
   strat
   # Remove row percents from all categorical variables (list is recycled)
   cat.rmstat = list('row'),
   # Remove missing, min/max from bili and no summary stats from copper
   cont.rmstat = list(c('miss', 'minmax'), 'None'),
   colnames = c('Variable', 'D-penicillamine', 'Placebo', 'Overall', 'p-value'),
   output
              = 'latex'
   )
```

out.plain

Plain Text Output

Description

This function outputs the results of calls to make.table, quick.table, cat.var, and cont.var to the console as formatted plain text data frames.

Usage

```
out.plain(tab, colnames)
```

Arguments

A single call to either cat.var or cont.var, or a data frame produced via

make.table or quick.table. Required.

colnames A vector of column names for the table output. Optional.

Details

A call to this function is the default output argument to all package functions that produce summary output.

quick.table 13

Value

A plain text data frame of summary statistics outputted to the console.

Author(s)

Erica Wozniak

See Also

Table1 make.table quick.table cat.var cont.var stat.col out.latex out.html

Examples

```
library(survival)
make.table(dat = pbc,
    cat.varlist = c('stage', 'sex'),
    cat.header = c('Stage', 'Sex'),
    cat.rownames = list(c('I', 'II', 'III', 'IV'), c('Male', 'Female')),
    cat.ptype = c('fisher', 'fisher'),
cont.varlist = c('bili', 'copper'),
    cont.header = c('Bilirubin', 'Urine copper'),
    cont.ptype = c('wilcox', 'wilcox'),
                 = c('trt'),
    # Remove row percents from all categorical variables (list is recycled)
    cat.rmstat = list('row'),
    \# Remove missing, min/max from bili and no summary stats from copper
    cont.rmstat = list(c('miss', 'minmax'), 'None'),
    colnames
                 = c('', 'D-penicillamine', 'Placebo', 'Overall', 'p-value')
    )
```

quick.table

Quick Summary Table

Description

Similarly to make.table, this function produces summary statistics for categorical and continuous variables and combines the results into a data frame for plain text, HTML, or LaTeX output. However, minimal input (a data frame only) is required to construct the table, as this function dynamically classifies variables as either categorical or continuous. This function is ideal for rapid summaries of data frames, but make.table should be used instead for custom labeling and significance testing.

Usage

```
quick.table(dat, strat, dec, colnames, output, classlim, stripe, stripe.col)
```

14 quick.table

Arguments

dat	Data frame input of all variables used for the summary table. Required.
strat	Character vector of names of one or more stratifying variables. Optional; defaults to no stratification.
dec	Integer value for number of decimal places to be printed for each summary statistic. Optional; defaults to 2.
colnames	Ordered vector of column names for table. Optional; defaults to levels of strata variable(s).
output	Choice of formatted table output: 'html', 'latex', or 'plain'. Optional; defaults to plain text printed in console.
classlim	Integer classification threshold for categorical and continuous variables. Optional; defaults to numeric variables with 6 or more levels being classified as continuous, and categorical otherwise. All non-numeric variables are classified as categorical.
stripe	HTML output only: logical indicator for whether to zebra stripe every other variable in the table. Optional; defaults to TRUE.
stripe.col	HTML output only: hex color code to be used for zebra striping the table. Optional; defaults to a light gray, #F7F7F7.

Value

Returns a table of summary statistics for specified categorical and continuous variables in the requested output format. The plain text default returns a data frame to the console.

Note

A few miscellaneous notes:

Table output to HTML and LaTeX require the package dependencies htmlTable and xtable

By default, the final summary table is rearranged such that the variables in the table are in the same order as in the raw data set (dat=).

If missing values occur in stratifying variables (strat=), then the observations with missings will be removed from the table, and a comment will be printed indicating the total number removed. If multiple stratifying variables are provided, look to the console for a summary of total missing observations for each variable.

For additional labeling and significance testing options, see make. table.

Author(s)

Erica Wozniak

See Also

Table1 make.table cat.var cont.var stat.col out.plain out.latex out.html

stat.col 15

Examples

```
library(survival)
attach(pbc)
##Plain text example
quick.table(dat=pbc,
    strat='sex',
                ', 'Male', 'Female', 'Overall')
    colnames=c('
##Example with HTML options
library(htmlTable)
quick.table(dat=pbc,
    strat='sex',
    colnames=c('', 'Male', 'Female', 'Overall'),
    #See www.color-hex.com for codes
    output='html',
    stripe.col='#d0f2b0'
    )
```

stat.col

Statistical Testing

Description

This function produces vectors of p-values and tests for addition via ptype= to a summary table generated with make.table. P-values are rounded to three decimal places or formatted as '<0.001' as appropriate. The name of the test is printed if requested with pname=TRUE.

Usage

```
stat.col(var, strat, ptype, pname, output)
```

Arguments

Variable input. Required. var Column stratifying variable. Required for two-sample tests. strat Test requested to compute p-value. Required. ptype 't.oneway' One-sample t-test 'wilcox.oneway' One-sample test for median=0 'prop.oneway' Binomial test for equal proportions (prop=0.5) 'chisq' Chi-square test 'fisher' Fisher's exact test 'ttest' t-test 'anova' One-way ANOVA 'kruskal' Kruskal-Wallis test 'wilcox' Wilcoxon rank-sum test 'ttest.pair' Paired t-test 'wilcox.pair' Wilcoxon signed-rank test

16 stat.col

'mcnemar' McNemar's test

pname If TRUE then the name of the statistical test used will be printed below the

p-value. Optional; defaults to FALSE.

output Type of output, passed from make.table or quick.table. Optional; defaults

to "plain".

Value

Returns a vector with a p-value and the name of the test used, if requested, to be added to a summary data frame.

Author(s)

Erica Wozniak

See Also

Table1 make.table quick.table cat.var cont.var out.plain out.latex out.html

Examples

```
library(survival)
attach(pbc)
##Default output
stat.col(var=bili,
         strat=sex,
         ptype='ttest')
##Print test
stat.col(var=bili,
         strat=sex,
         ptype='ttest',
         pname=TRUE)
##The ptype= and pname= options in summary functions call stat.col
#Results are automatically added to the summary table
cat.var(var=hepato,
        strat=sex,
        rownames=c('Not present', 'Present'),
        header=c('Hepatomegaly'),
        ptype='chisq',
        pname=TRUE
```

Index

```
*Topic package
    Table1-package, 2
cat.var, 3, 3, 6, 8, 10–14, 16
cont.var, 3, 4, 5, 6, 8, 10-14, 16
htmlTable, 2, 8, 9, 14
make.table, 2-6, 6, 9-16
out.html, 3, 4, 6, 8, 9, 12-14, 16
out.latex, 3, 4, 6, 8, 10, 11, 13, 14, 16
out.plain, 3, 4, 6, 8, 10, 12, 12, 14, 16
quick.table, 2-4, 6, 8-13, 13, 16
rbind, 11
stat.col, 3–8, 10, 12–14, 15
Table1, 4, 6, 8, 10, 12–14, 16
Table1 (Table1-package), 2
Table1-package, 2
xtable, 2, 8, 11, 14
```