gtsummary

Daniel D. Sjoberg, Karissa Whiting, Margaret Hannum, Emily C. Zabor, Michael Curry

Introduction

Data Summaries

Table 1: Example data frame, `trial`

colname	label	class	values
trt	Chemotherapy Treatment	character	`Drug A`, `Drug B`
age	Age	numeric	`6`, `9`, `10`, `17` , ...
marker	Marker Level (ng/mL)	numeric	`0.003`, `0.005`, `0.013`, `0.015` , ...
stage	T Stage	factor	`T1`, `T2`, `T3`, `T4`
grade	Grade	factor	`I`, `II`, `III`
response	Tumor Response	integer	`0`, `1`
death	Patient Died	integer	`0`, `1`
ttdeath	Months to Death/Censor	numeric	`3.53`, `5.33`, `6.32`, `7.27` , ...

`tbl_summary()`

```
tbl_summary_1 <-
  trial %>%
  select(age, grade, response, trt) %>%
  tbl_summary(by = trt)
```

Table 2: `tbl_summary()` function arguments

Argument	Description
`label`	specify the variable labels printed in table
`type`	specify the variable type (e.g. continuous, categorical, etc.)
`statistic`	change the summary statistics presented
`digits`	number of digits the summary statistics will be rounded to
`missing`	whether to display a row with the number of missing observations
`missing_text`	text label for the missing number row
`sort`	change the sorting of categorical levels by frequency
`percent`	print column, row, or cell percentages

```
tbl_summary_2 <-
  trial %>%
```

```
  select(age, grade, response, trt) %>%
  tbl_summary(by = trt,
              label = age ~ "Patient Age",
              statistic = list(all_continuous() ~ "{mean} ({sd})",
                               all_categorical() ~ "{n} / {N} ({p}%)"),
              digits = list(age ~ 2,
                            all_categorical() ~ c(0, 0, 1)),
              missing = "no")
```

Table 3: `tbl_summary()` function arguments

Function	Description
`add_p()`	add p-values to the output comparing values across groups
`add_overall()`	add a column with overall summary statistics
`add_n()`	add a column with N (or N missing) for each variable
`add_difference()`	add column for difference between two group, confidemce interval, and p-value
`add_stat_label()`	add label for the summary statistics shown in each row
`add_stat()`	generic function to add a column with user-defined values
`add_q()`	add a column of q values to control for multiple comparisons

Figure 1: Simple ‘tbl_summary()‘ example

Characteristic	**Drug A**, N = 98[1]	**Drug B**, N = 102[1]
Age	46 (37, 59)	48 (39, 56)
Unknown	7	4
Grade		
I	35 (36%)	33 (32%)
II	32 (33%)	36 (35%)
III	31 (32%)	33 (32%)
Tumor Response	28 (29%)	33 (34%)
Unknown	3	4

[1] Median (IQR); n (%)

`tbl_svysummary()`

`tbl_cross()`

`tbl_survfit()`

Customization

Model Summaries

`tbl_regression()`

`tbl_uvregression()`

In-line Reporting

Merging and Stacking

Themes

Print Engines