

Summary Table – Univariate Association with a Numerical Outcome

Macro Name: UNI_NUM

Created Date/Author: Feb. 2012/Yuan Liu

Last Update Date/Person: Oct. 3, 2013/Dana Nickleach

Current Version: V9

Working Environment: SAS 9.3 English version

Contact: Dr. Yuan Liu yliu31@emory.edu

Purpose: To produce a summary table of the univariate association of a numerical outcome with all other covariates in the data individually. For categorical covariates, the sample size, mean and median along with ANOVA test (parametric p-value) or Kruskal-Wallis test (non-parametric p-value) will be carried out. For numerical covariates, both Pearson correlation coefficient (parametric p-value) and Spearman correlation coefficient (non-parametric p-value) will be calculated.

Notes: 1) The order of variables in the summary table is the same as the input order. For the best results, you may want to put the demographic variables together and also clinical characteristics variables together; 2) The biostatistician may need to help investigator decide which statistics (parametric or non-parametric p-value) is more appropriate to the data.

Parameters:

Macro variable	Description
DATASET	The name of the data set to be analyzed.
OUTCOME	The numerical outcome variable.
CLIST	List of categorical variables, separated by empty space.
NLIST	List of numerical variables, separated by empty space.
NONPAR	Set to F to suppress non-parametric statistics. The default value is T.
OUTPATH	Path for output table to be stored.
FNAME	File name for output table.
DEBUG	Set to T if running in debug mode (optional). Work datasets will not be deleted in debug mode. This is useful if you are editing the code or want to further manipulate the resulting data sets. The default value is F.

Usage Example:

```
DATA analysis;
  input id os_censor Sex $ Age duration os progress $ trt $;
  LABEL os = 'Overall Survival (months)'
        progress = 'Progression'
        trt = 'Treatment'
        duration = 'Duration of Radiation';
  DATALINES;
1      1      M      40      44      20      No      B
2      1      F      45      46      16      Yes     A
3      1      F      40      32      20      No      B
4      1      F      47      32      23      No      B
5      0      M      41      25      22      No      B
6      1      M      54      35      13      No      B
7      1      M      48      50      9       Yes     A
8      1      M      36      33      12      Yes     B
9      0      F      49      51      8       Yes     A
10     1      M      49      52      10      Yes     A
11     1      M      44      35      12      No      A
12     1      M      49      50      8       Yes     A
13     1      M      44      44      14      Yes     A
14     1      M      50      31      10      Yes     A
15     1      M      53      40      15      No      B
16     0      M      52      29      20      Yes     B
17     1      F      46      45      5       Yes     A
18     1      F      37      44      11      Yes     A
19     1      M      49      46      13      No      B
20     1      M      42      31      11      No      A
;

TITLE 'Table 2 Univariate Association with Duration';
%UNI_NUM(dataset = analysis,
  outcome = duration,
  clist = trt sex,
  nlist = age,
  outpath = C:\Documents and Settings\User\My Documents\,
  fname = Table 2 Univariate Association with Duration);
TITLE;
```

Summary Table Example:

1

Table 2 Univariate Association with Duration						
Duration of Radiation						
Variable	Level	N	Mean	Median	ANOVA P-value	Kruskal-Wallis P-value
Treatment	A	11	43.55	45.00	0.021	0.036
	B	9	35.11	33.00		
Sex	F	6	41.67	44.50	0.519	0.482
	M	14	38.93	37.50		

Duration of Radiation					
Variable	N	Pearson CC	Pearson P-value	Spearman CC	Spearman P-value
Age	20	0.183	0.441	0.177	0.456

Permission:

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Log of Updates:

Date	By	Description	Version
8/3/2012	Dana Nickleach (dnickle@emory.edu)	Added deletion of temporary data sets created by the macro, restoration of original options at the end of the macro, and added debug parameter.	V2
8/7/2012	Dana Nickleach (dnickle@emory.edu)	Prevented analysis data set from being overwritten in PROC SORT and prevented error if both categorical and numerical variables are not supplied. Fixed capitalization sensitivity in outcome variable.	V3
10/16/2012	Sungjin Kim (skim61@emory.edu)	Added check for outcome variables that are too long. Got rid of leading space in RTF file name. Changed lengths of 100 to 96 to avoid warnings in proc report.	V4
10/23/12	Dana Nickleach (dnickle@emory.edu)	Fixed proc report ODS listing error, changed “group” to “order” to prevent note, added spanrows option, changed length of covariate and outcome labels to max label length to avoid truncation, and prevented case sensitivity for T/F parameters.	V5
12/21/12	Dana Nickleach (dnickle@emory.edu)	Increased output format for correlation coefficients to sufficient width, fixed bug preventing character values of “-” from printing in the table, changed column header to more accurately say “Kruskal-Wallis”, and removed the page break between tables.	V6
3/11/13	Dana Nickleach (dnickle@emory.edu)	Added NONPAR parameter.	V7
5/9/13	Dana Nickleach (dnickle@emory.edu)	Added check for CLIST variables with less than two levels, otherwise incorrect p-values will be printed.	V8
10/3/13	Dana Nickleach (dnickle@emory.edu)	Changed BY statement to CLASS statement in PROC MEANS to avoid splitting up formatted categories.	V9