THE TABLES PROCEDURE AS AN ALTERNATIVE TO THE DATA _NULL_ OR PRINT PROCEDURE IN GENERATING REPORT-QUALITY STATISTICAL TABLES

PATRICIA L. GEREND, SYNTAX

ABSTRACT

PROC TABLES is a user-written, user-supported procedure that computes descriptive statistics, tests control against experimental groups, and generates a report-quality table of the statistics and test results.

When compared to the use of PROCS MEANS or GLM combined with a PROC PRINT or a DATA _NULL_, PROC TABLES had these advantages: everything is done in one step; table is report quality; labeling columns is easier; statements are simpler; statistically-significant differences are automatically marked in the table. PROC TABLES had these disadvantages: footnotes, user-defined formats, and the missing option do not function; triple-spacing is mandatory; the procedure is not validated by SAS® Institute.

INTRODUCTION

While generating descriptive statistics is typically a very easy task in SAS®, producing a report-quality table of these statistics can be quite time-consuming. The DATA _NULL_ is the SAS® tool that is most effective at producing report quality tables, but it is also the most difficult and tedious. PROC PRINT can easily generate a table, but the report quality is marginal.

This paper will discuss the use of PROC TABLES as an alternative to the use of PROC MEANS (for descriptive statistics) or PROC GLM (for statistics and testing) in combination with a PROC PRINT or DATA _NULL_ for printing. PROC TABLES can accomplish all of these simply, quickly, and in a quality fashion. The user must keep in mind, however, that the table format has limited flexibility and that the procedure is not validated by SAS® Institute.

METHODS / DISCUSSION

PROC TABLES is capable of doing the following: computing descriptive statistics, testing control against experimental groups, and producing a table of the statistics and test results.

The descriptive statistics that can be requested are mean, percent, standard deviation and error, number, count, number missing, sum, maximum, minimum, range, variance, uncorrected and corrected sums of squares, and coefficient of variation. These statistical capabilities are reasonably similar to those of PROCS MEANS and SUMMARY. In order to produce these statistics, the data must be sorted on the CLASS variable.

In addition to generating such statistics, PROC TABLES also offers statistical comparison options. The user can choose between several statistics (Bonferroni, Williams, Dunnet, Student), define the significance level, request a one- or two-tailed test, request computations on ranked or raw data, and specify a control group for comparison with experimental groups. These test capabilities compare to those available with PROC GLM.

The most significant time-saving feature of PROC TABLES is the fact that a quality table is generated directly from the procedure; no added printing procedure or data step is necessary. There are a couple of non-standard format options available: STACK and FORMAT. Through the STACK option, up to 10 lines of variable labels can be specified (as opposed to 3 lines with PROC PRINT); the LABEL statement used with this option must contain a delimiter, defined as the first character of the label, to split the label into a stack. The FORMAT option allows the user to link the number of significant digits for each variable or for the entire table to the average standard deviation for each variable or for the entire table.

While an acceptable table is very easy to produce with PROC TABLES, with ease comes a certain amount of inflexibility. It is not possible to use some of the standard options, such as footnotes, user-defined formats, and the missing option. It is also not possible to control the spacing; triple-spacing is automatic. Text-editting of the table would be required to alter any of these situations.

The code below was used on pharmaceutical data to produce the attached table.

PROC SORT;
  BY TREATMENT;
PROC TABLES
  FORMAT = T2  /* no. sig. digits */
  STACK  /* labels */
  MEAN STD  /* desc. stat. */
  STATISTIC = BONFERRONI
  LEVEL = 0.05
  CONTROL = 'PLACEBO';
CLASS TREATMENT;
VAR DAY1 DAY2 DAY3 DAY4;
TITLE 'METABOLITE LEVELS';
OPTIONS MODATE NONUMBER;
Using PROC PRINT in combination with a statistical PROC to generate a report table would have the advantages of standard options, but would require text-editing of the labels to position them over both the mean and standard deviation variables or to achieve spacing other than single or double. In addition, the final product would not be as readable as that of PROC TABLES.

A combination of the DATA _NULL_ and a statistical procedure offers the most flexibility in that a variable, label, title, or footnote can be placed anywhere on the page. The price paid for this flexibility, however, is the programmer time involved in specifying exactly where to put all the information.

**CONCLUSION**

PROC TABLES is a reasonable alternative to the use of a statistical procedure in combination with a PROC PRINT or a DATA _NULL_ in that it is a one-step method, the program statements are simple, very little programmer time is required, and the table is report quality. The drawbacks include the fact that the procedure is not validated by SAS® Institute and that some standard options do not function.

**METABOLITE LEVELS**

<table>
<thead>
<tr>
<th>TREATMENT</th>
<th>STATISTIC</th>
<th>DAY 1 POST-MEAL</th>
<th>DAY 2 PRE-MEAL</th>
<th>DAY 3 POST-MEAL</th>
<th>DAY 4 PRE-MEAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>MEAN</td>
<td>102</td>
<td>103</td>
<td>102</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>PLACEBO</td>
<td>MEAN</td>
<td>91</td>
<td>89</td>
<td>85</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>X</td>
<td>MEAN</td>
<td></td>
<td>101</td>
<td>104</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td></td>
<td>10</td>
<td>5</td>
<td>103</td>
</tr>
<tr>
<td>Y</td>
<td>MEAN</td>
<td>107</td>
<td>108</td>
<td>117</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>19</td>
<td>10</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>Z</td>
<td>MEAN</td>
<td>101</td>
<td>98</td>
<td>103</td>
<td>115*</td>
</tr>
<tr>
<td></td>
<td>STD</td>
<td>16</td>
<td>20</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>

* : P <= .05, TWO TAILED BONFERRONI T ON RAW DATA.
@ : TREATED AS CONTROL GROUP FOR STATISTICAL TESTING.