Paper 3253-2015

Need Additional Statistics in Your Report? - ODS OUTPUT to the Rescue!

Deborah Buck, inVentiv Health

.

ABSTRACT

You might be familiar with or experienced in writing or running reports using PROC REPORT, PROC TABULATE, or other methods of report generation. These reporting methods are often very flexible, but they can be limited in the statistics that are available as options for inclusion in the resulting output.

SAS® provides the capability to produce a variety of statistics through Base SAS® and SAS/STAT® procedures by using ODS OUTPUT objects. These procedures include statistics from PROC CORR, PROC FREQ, and PROC UNIVARIATE in Base SAS, as well as PROC GLM, PROC LIFETEST, PROC MIXED, PROC LOGISTIC, and PROC TTEST in SAS/STAT. A number of other procedures can also produce useful ODS OUTPUT objects.

Commonly requested statistics for reports include p-values, confidence intervals, and test statistics. These values can be computed with the appropriate procedure and then use ODS OUTPUT to output the desired information to a data set and include the new information with the other data used to produce the report.

Examples that demonstrate how to easily generate the desired statistic or other information and include it to produce the requested final reports are provided and discussed.

INTRODUCTION

For many years the statistics or information produced by a SAS procedure (PROC) could be easily seen within the PROC output, but often were very difficult to extract for use in other reports or applications. Some PROCs, such as MEANS, SUMMARY, FREQ, and REPORT allowed outputting of datasets with selected statistics using the OUT= option or OUTPUT= statement, but results of all statistical tests were not always readily available. These methods were dependent upon the PROC being used. Another approach to extract the needed information if it was not available from an output dataset was to use PROC PRINTTO and to tell SAS exactly what line and column to go to for the information and the width/informat of the field to read. Needless to say, this was a very time-consuming, tedious process and involved quite a bit of 'hit or miss' programming.

When ODS (Output Delivery System) programming became available in SAS a new, standardized way of outputting information from PROCs also became available where procedures utilize the same straight-forward approach to obtain output from almost every SAS procedure - ODS OUTPUT objects.

ODS TRACE STATEMENT

Where or how can you learn which PROCs have objects available to use? You can refer to the documentation for each version of SAS which has a section under SAS 9.x Output Delivery System: User's Guide that lists PROCs, the object table names for each, the description and any necessary options. These are also grouped into 3 categories - Base SAS Procedures, SAS STAT Procedures and SAS ETS Procedures - for ease in finding the desired information.

However, it is probably easier to use the ODS TRACE statement. The ODS TRACE statement will provide the name of each output object produced by a given PROC.

The form of the ODS TRACE statement is:

```
ods trace on / label;
  < SAS procedure statements >
ods trace off;
```

For the < *SAS procedure statements* > portion above you include the PROC statement in the form that you would program to see the information in the normal PROC output.

EXAMPLE 1 – FISHER'S EXACT TEST

This example is based on a small pilot study comparing a new treatment with a traditional treatment on change in a laboratory parameter. A total of 30 subjects were enrolled with 15 subjects on each treatment. The variables AGE, GENDER, METHOD and CHANGE are included in SAS dataset METHOD2. The first report desired generates the cross-tab of GENDER by METHOD with Fisher's Exact Test to check that the proportions of male and female are comparable on each treatment (METHOD).

The following is the desired output from PROC REPORT:

Gender by Method Output with Fisher's Exact Test Probability Value

Gender	New Treatment	Traditional Treatment	Probability
Female	10 (66.67%)	7 (46.67%)	0.462
Male	5 (33.33%)	8 (53.33%)	

Output 1. Output of Gender by Method Report from PROC REPORT

In order to produce the above table several steps are necessary. These include:

1. Using the ODS TRACE statement to determine the objects which contain the frequencies, percentages and appropriate Fisher's Exact Test p-value desired.
2. Creating SAS datasets from the ODS OUTPUT objects and manipulating the data to generate the variables in the required form.
3. Transposing the data to be in the correct form for PROC REPORT to produce a detail report.
4. Generating the PROC REPORT output.

The ODS TRACE is included in the SAS program in the following form:

```
ods trace on/label listing;    *ODS TRACE turned on in output window;
 proc freq data=method2;
  tables gender*method/exact;  *requested PROC FREQ table and statistics;
 title 'ODS TRACE on for PROC FREQ';
 run;
ods trace off;                 *ODS TRACE turned off;
```

This code segment tells SAS to apply ODS TRACE and write the output and the ODS TRACE information to the listing. (By default, this information is written to the SAS log.) Note in the sections below the yellow highlighted lines show the name of each ODS OUTPUT object created in this PROC step. This NAME: value will be used in the next PROC step to create and name the new datasets.

The ODS OUTPUT object information for the frequencies and probabilities is shown below.

```
                                                ODS TRACE on for PROC FREQ
                                                    The FREQ Procedure
Output Added:
-------------
Name:        CrossTabFreqs
Label:       Cross-Tabular Freq Table
Template:    Base.Freq.CrossTabFreqs
Path:        Freq.Table1.CrossTabFreqs
Label Path: 'The Freq Procedure'.'Table gender * method'.'Cross-Tabular Freq Table'
-------------
                                                     Table of gender by method

                                              gender(Gender)     method(Treatment)

                                              Frequency|
                                              Percent  |
                                              Row Pct  |
                                              Col Pct  |NEW     |OLD     |  Total
                                              ---------+--------+--------+
                                              F        |     10 |      7 |     17
                                                       |  33.33 |  23.33 |  56.67
                                                       |  58.82 |  41.18 |
                                                       |  66.67 |  46.67 |
                                              ---------+--------+--------+
                                              M        |      5 |      8 |     13
                                                       |  16.67 |  26.67 |  43.33
                                                       |  38.46 |  61.54 |
                                                       |  33.33 |  53.33 |
                                              ---------+--------+--------+
                                              Total          15       15       30
                                                             50.00    50.00   100.00
```

Output 2. First section of output from ODS TRACE on PROC FREQ.

The following section of the PROC FREQ output includes information on Chi-Square statistics.

```
                                                ODS TRACE on for PROC FREQ
                                                    The FREQ Procedure


                                             Statistics for Table of gender by method
Output Added:
-------------
Name:        ChiSq
Label:       Chi-Square Tests
Template:    Base.Freq.ChiSq
Path:        Freq.Table1.ChiSq
Label Path: 'The Freq Procedure'.'Table gender * method'.'Chi-Square Tests'
-------------
                                    Statistic                     DF       Value      Prob
                                    ------------------------------------------------------
                                    Chi-Square                     1      1.2217    0.2690
                                    Likelihood Ratio Chi-Square    1      1.2308    0.2673
                                    Continuity Adj. Chi-Square     1      0.5430    0.4612
                                    Mantel-Haenszel Chi-Square     1      1.1810    0.2772
                                    Phi Coefficient                       0.2018
                                    Contingency Coefficient               0.1978
                                    Cramer's V                            0.2018
```

Output 3. Second section of output from ODS TRACE on PROC FREQ.

The third segment below shows the information relating to the Fisher's Exact test.

```
Output Added:
-------------
Name:        FishersExact
Label:       Fisher's Exact Test
Template:    Base.Freq.ChisqExactFactoid
Path:        Freq.Table1.FishersExact
Label Path: 'The Freq Procedure'.'Table gender * method'.'Fisher''s Exact Test'
-------------

                                                ODS TRACE on for PROC FREQ

                                                    The FREQ Procedure

                                            Statistics for Table of gender by method

                                                    Fisher's Exact Test
                                               ---------------------------------
                                               Cell (1,1) Frequency (F)        10
                                               Left-sided Pr <= F          0.9303
                                               Right-sided Pr >= F         0.2311

                                               Table Probability (P)       0.1614
                                               Two-sided Pr <= P           0.4621

                                                     Sample Size = 30
```

Output 4. Third section of output with ODS TRACE on PROC FREQ.

You now know the names of the ODS OUTPUT objects that are needed to create corresponding SAS datasets and can apply this in the next segment of code.

```
ods output crosstabfreqs=cross;  *Output dataset for crosstabs;
ods output fishersexact=fishers; *Output dataset for Fisher's Exact test;
 proc freq data=method2;
  tables gender*method/exact;
 title 'Outputting 2 datasets with PROC FREQ information';
 run;
```

The above code creates two SAS datasets - CROSS and FISHERS - which contain the information that is needed for the desired report. However, it does not show the variable names, types or values. In order to explore these PROC PRINT and PROC CONTENTS are used.

The output from PROC PRINT is shown is the following Output 5 below. From this you can see that the variables gender, method, frequency and colpercent are the desired fields and that only _TYPE_=11 (combinations of the variables GENDER and METHOD) should be included (highlighted in yellow).

Printing of CROSS dataset with Frequencies and Percentages

Obs	Table	gender	method	_TYPE_	_TABLE_	Frequency	Percent	Row Percent	Col Percent	Missing
1	Table gender * method	F	NEW	11	1	10	33.333	58.8235	66.6667	.
2	Table gender * method	F	OLD	11	1	7	23.333	41.1765	46.6667	.
3	Table gender * method	F		10	1	17	56.667	.	.	.
4	Table gender * method	M	NEW	11	1	5	16.667	38.4615	33.3333	.
5	Table gender * method	M	OLD	11	1	8	26.667	61.5385	53.3333	.
6	Table gender * method	M		10	1	13	43.333	.	.	.

7	Table gender * method	NEW	01	1	15	50.000	.	.	.
8	Table gender * method	OLD	01	1	15	50.000	.	.	.
9	Table gender * method		00	1	30	100.000	.	.	0

Output 5. Output of WORK.CROSS from PROC PRINT.

Information from PROC CONTENTS is used to capture the attribute information for variables in CROSS.

Similarly, PROC PRINT of FISHERS dataset results in the following output below and shows that the desired observation is Name1='XP2_FISH' and the variable cValue1 contains the probability value to be included in the final report.

Printing of FISHERS dataset

Obs	Table	Name1	Label1	c Value1	nValue1
1	Table gender * method	Cell1_FREQ	Cell (1,1) Frequency (F)	10	10.000000
2	Table gender * method	XPL_FISH	Left-sided Pr <= F	0.9303	0.930290
3	Table gender * method	XPR_FISH	Right-sided Pr >= F	0.2311	0.231069
4	Table gender * method				.
5	Table gender * method	P_TABLE	Table Probability (P)	0.1614	0.161359
6	Table gender * method	XP2_FISH	Two-sided Pr <= P	0.4621	0.462137

Output 6. Output of WORK.FISHERS from PROC PRINT.

The following code segment manipulates the variables into the desired character variables.

```
*Get frequency and column percent information from PROC FREQ;
   data cross1;
     set cross;
      if _type_='11';
   keep gender method frequency colpercent;
   run;

*Modify N and Percent to create character formats for PROC REPORT;
   data cross2;
    set cross1;

    length n_perc $12;

    n=put(frequency,3.);
    perc=put(colpercent, 5.2);
    n_perc=(n||' ('||perc||'%'||')');  *Variable in form xx (xx.xx%);
   run;

   proc sort data=cross2;
    by gender method;
   run;
```

This is followed by PROC TRANSPOSE to change the orientation so that there is one record for each GENDER that includes both METHOD values.

```
*Switch orientation;
   proc transpose data=cross2 out=newstats1;
    by gender ;
     id method;
     var n_perc;
   run;
```

The NEWSTATS1 SAS dataset is then merged with the FISHERS dataset to produce the final dataset.

```
*Include probability from PROC TTEST;
   data cross_fish;
    merge newstats1 fishers(where=(name1='XP2_FISH') keep=name1 cvalue1);

     prob=put(cvalue1,5.3);
   drop name1 cvalue1;
   run;
```

Finally, the following PROC REPORT step produces the desired crosstab report with frequencies, column percentages and the p-value associated with the Fisher's Exact test. (Shown in Output 1.)

```
*Final report for Gender by Treatment;
   proc report data=cross_fish nowd headline headskip;
    column gender new old prob;
     define gender/'Gender' display left width=12 format=$gfmt.;
     define new/'New Treatment' display center width=14;
     define old/'Traditional Treatment' display center width=14;
     define prob/'Probability' display center width=12;
   title "Gender by Treatment Output with Fisher's Exact Test Probability Value";
   run;
```

EXAMPLE 2 – UNPAIRED T-TEST

On this same study there is a report required showing descriptive statistics on change from baseline by treatment and overall which includes the number, mean, standard deviation, median, minimum, maximum and probability from the unpaired t-test of the means of the treatments. The following shows the required output.

PROC REPORT EXAMPLE with STATISTICS FROM ODS OUTPUT for Change by Method

Statistic	New Treatment	Traditional Treatment	All
N	15	15	30
Mean (SD)	2.67 (1.99)	0.33 (2.47)	1.50 (2.50)
Median	3.0	1.0	2.0
Minimum-Maximum	(-1.0-6.0)	(-5.0-4.0)	(-5.0-6.0)
p-value from t-test for difference between Method means			0.008

Output 7. Final PROC REPORT output for CHANGE by METHOD of Treatment.

Again, it's ODS OUTPUT objects to the rescue to provide the desired information.

For this table you can use PROC MEANS and PROC TTEST to obtain the desired statistics which can then be modified to the format desired in the PROC REPORT generation.

First, the ODS TRACE ON is used with the PROC MEANS and PROC TTEST to produce the following messages in the SAS LOG:

ODS TRACE ON object information from PROC MEANS

```
Output Added:
-------------
Name:        Summary
Label:       Summary statistics
Template:    base.summary
Path:        Means.Summary
```

```
Label Path: 'The Means Procedure'.'Summary statistics'
```

Output 8. Output from ODS TRACE ON on summary variables from PROC MEANS.

ODS TRACE ON object information from PROC TTEST

```
Output Added:
-------------
Name:       Statistics
Label:      Statistics
Template:   Stat.TTest.Statistics
Path:       Ttest.change.Statistics
Label Path: 'The Ttest Procedure'.'Change in Lab Test'.'Statistics'
-------------

Output Added:
-------------
Name:       ConfLimits
Label:      Confidence Limits
Template:   Stat.TTest.ConfLimits
Path:       Ttest.change.ConfLimits
Label Path: 'The Ttest Procedure'.'Change in Lab Test'.'Confidence Limits'
-------------

Output Added:
-------------
Name:       TTests
Label:      T-Tests
Template:   Stat.TTest.TTests
Path:       Ttest.change.TTests
Label Path: 'The Ttest Procedure'.'Change in Lab Test'.'T-Tests'
-------------

Output Added:
-------------
Name:       Equality
Label:      Equality of Variances
Template:   Stat.TTest.Equality
Path:       Ttest.change.Equality
Label Path: 'The Ttest Procedure'.'Change in Lab Test'.'Equality of Variances'
-------------
```

Output 9. Output from PROC TTEST showing all ODS OUTPUT Objects created.

As you can see there is one object generated by PROC MEANS containing all descriptive statistics and four objects produced by PROC TTEST covering the statistics, confidence limits, ttest results and equality of variance tests.

The above information is included with the PROCs to generate both outputs and create new SAS datasets using the following SAS code.

```
*Output dataset with overall and by method descriptive statistics from  PROC
MEANS;
   ods output summary=work.summary1;
    proc means data=method2 n mean stddev median min max;
     ways 0 1;        *Derives statistics overall and by method;
     class method;
     var change;
    title 'Descriptive Statistics for Change by Method';
```

```
   run;


*Output ttest probability from PROC TTEST;
   ods output ttests=work.ttest;
    proc ttest data=method2;
     class method;
     var change;
    title 't-test for Change by Method';
   run;
```

Printing of the SUMMARY1 dataset from the ODS OUTPUT object produces the following output. Note that the WAYS 0 1 statement in PROC MEANS provides summary statistics both overall and by method. The value of METHOD for the overall statistics is missing because it includes both NEW and OLD.

Printing of SUMMARY1 dataset

Obs	method	NObs	change_N	change_Mean	change_ StdDev	change_ Median	change_Min	change_Max
1		30	30	1.5	2.5017235438	2	-5	6
2	NEW	15	15	2.6666666667	1.9880595948	3	-1	6
3	OLD	15	15	0.3333333333	2.4688535994	1	-5	4

Output 10. Output of SUMMARY1 dataset from PROC PRINT.

The PROC PRINT of the new TTEST dataset generates the following output. From the PROC TTEST output you see that the variances are equal and therefore choose to use the probability associated with the t-test using the Method=Pooled.

Print of t-test dataset

Obs	Variable	Method	Variances	tValue	DF	Probt
1	change	Pooled	Equal	2.85	28	0.0081
2	change	Satterthwaite	Unequal	2.85	26.782	0.0083

Output 11. Output of TTEST dataset from PROC PRINT.

The following code modifies the descriptive statistics to the desired forms and merges that information with the probability value from the t-test.

```
*Create character data from numeric results and format into form for
 PROC REPORT;
   data summary2;
    set summary1 (in=a) ;

     if method=' ' then method='ALL'; *Assign ALL value to overall column;

      n=put(change_n,3.);
      mean=put(change_mean,5.2);
      sd='('||put(change_stddev,5.2)||')';
      mean_sd=mean || ' ' ||compress(sd);
      med=put(change_median,5.1);
      min=put(change_min,5.1);
      max=put(change_max,5.1);

      min_max='('||(compress(min)||'-'||compress(max))||')';

    keep method n mean_sd med min_max;
   run;
```

```
*Create dataset with only desired t-test statistics;
   data ttesta;
    set ttest(where=(method='Pooled'));

     method='ALL';   *Want the probability to be displayed in 'All' column;
     prob=put(probt,5.3);

    keep method prob;
   run;

*Merge descriptive statistics with t-test statistic;
   data summary2a;
    merge summary2(in=a) ttesta(in=b);
     by method;
   run;
```

The final step is to transpose the data which results in the following transposed dataset.

Printing of new transposed NEWSTATS Dataset

Obs	_NAME_	ALL	NEW	OLD
1	n	30	15	15
2	mean_sd	1.50 (2.50)	2.67 (1.99)	0.33 (2.47)
3	med	2.0	3.0	1.0
4	min_max	(-5.0-6.0)	(-1.0-6.0)	(-5.0-4.0)
5	prob	0.008		

Output 12. Output of transposed dataset NEWSTATS.

The following PROC FORMAT and PROC REPORT steps produce a detail report with the desired information. (Results shown in Output 7.)

```
   proc format;
    value $statfmt
          'n'='N'
          'mean_sd'='Mean (SD)'
          'med'='Median'
          'min_max'='Minimum-Maximum'
          'prob'='p-value from t-test for difference between Method means'
    ;
   run;

*Final report for Change by Method statistics;
   proc report data=newstats nowd headline headskip split='^';
    column _name_ new old all;
     define _name_/'Statistic' display left width=36 format=$statfmt. flow;
     define new/'New Treatment' display center width=14;
     define old/'Traditional Treatment' display center width=14;
     define all/'All' display center;
    title1 'PROC REPORT EXAMPLE with STATISTICS FROM ODS OUTPUT for Change by Method';
   run;
```

EXAMPLE 3 - PROC COMPARE

The previous examples both focused on descriptive or inferential statistics. However, sometimes a report consists of additional types of information. For example, a report may be the result of PROC CONTENTS or PROC COMPARE output. ODS OUTPUT objects are created with these types of PROCs also.

An example of this output might be to show information on two datasets being compared for validation purposes including details on differences between variable values using PROC COMPARE. The following shows ODS TRACE ON with PROC COMPARE

```
*Example with PROC COMPARE;
   ods trace on/label ;
    proc compare base=method2 compare=replica;
     id id;
    title 'PROC COMPARE output';
    run;
   ods trace off;
```

This ODS TRACE ON produces the following output in the SAS log.

```
Output Added:
-------------
Name:        CompareDatasets
Label:       Datasets
Data Name:   BatchOutput
Path:        Compare.CompareDatasets
Label Path: 'The Compare Procedure'.'Datasets'
-------------

Output Added:
-------------
Name:        CompareSummary
Label:       Summary
Data Name:   BatchOutput
Path:        Compare.CompareSummary
Label Path: 'The Compare Procedure'.'Summary'
-------------

Output Added:
-------------
Name:        CompareDifferences
Label:       Differences
Data Name:   BatchOutput
Path:        Compare.CompareDifferences
Label Path: 'The Compare Procedure'.'Differences'
-------------
```

Output 13. Output of ODS TRACE in SAS LOG for PROC COMPARE.

Running the following code produces two datasets – CHECK1 and CHECK2.

```
   ods output comparedatasets=check1;
   ods output comparedifferences=check2;
```

```
 proc compare base=method2 compare=replica;
  id id;                  *Use the ID variable in the datasets for matching;
 title 'PROC COMPARE output';
run;
```

Printing of CHECK1 dataset provides the following output. Keeping the section on Data Set Summary in observations 7-10 for the variable 'batch' would show a quick summary of dataset names, creation and modification dates, number of variables and number of observations that could be included in the body of a report showing information about the compared datasets.

Information in Datasets Comparison - Output in CHECK1 Dataset

Obs	type	batch
1	h	The COMPARE Procedure
2	h	Comparison of WORK.METHOD2 with WORK.REPLICA
3	h	(Method=EXACT)
4	h	
5	h	Data Set Summary
6	h	
7	h	Dataset Created Modified NVar NObs
8	d	
9	d	WORK.METHOD2 31MAR15:20:52:40 31MAR15:20:52:40 6 30
10	d	WORK.REPLICA 01APR15:17:06:52 01APR15:17:06:52 6 30
11	d	
12	d	
13	h	Variables Summary
14	h	
15	d	Number of Variables in Common: 6.
16	d	Number of ID Variables: 1.

Output 14. Output from PROC PRINT of CompareDatasets object.

Similarly, the CHECK2 dataset from the CompareDifferences could be used in a report showing where any differences existed between the two datasets in the variable values.

CONCLUSION

This paper shows just a few examples of ODS OUTPUT objects and how easily they can be used in various types of reports. Although the descriptive and inferential statistics in these examples can be obtained multiple ways, creating a dataset from ODS OUTPUT objects can also make statistics from much more complex statistical procedures and models more readily available. ODS OUTPUT objects also provide a much easier way to extract text-type information from some PROCs for inclusion in other applications.

Hopefully, the next time you need statistics or information from a SAS procedure to use in another DATA step, PROC step or report ODS OUTPUT objects will come to mind!

REFERENCES AND RECOMMENDED READING

SAS Institute. 2012. SAS® 9.3 Output Delivery System: User's Guide, Second Edition. Cary, North Carolina: SAS Press.

Zender, Cynthia, 2011. "The Greatest Hits: ODS Essentials Every User Should Know." *Proceedings of the SAS Global Forum 2011 Conference.* Cary, NC: SAS Institute Inc. (Paper 300-2011)

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Deborah Buck
inVentiv Health
Houston, TX
281-723-3678
deborah.buck@inventivhealth.com