MINIMIZING MIXED OUTPUT

STEFANIE J. SILVA, LEWIN-TAG, INC.
DAVE P. MILLER, LEWIN-TAG, INC.

ABSTRACT

PROC MIXED is a powerful tool capable of producing numerous statistics; however, it is often the case that only two or three key statistics per model are required to summarize the results. When many models are run involving different response variables, subsets of cases, or minor variations in the model structure, thumbing through the numerous pages of output to locate statistics of primary interest becomes tedious. The Output Delivery System (ODS), a feature of PROC MIXED, used in conjunction with PROC DATASETS, PROC APPEND, the DATA step, and PROC PRINT can simplify this process. Our method uses macros to be easier to program. This use of ODS eliminates the need to print excessive amounts of paper and ultimately produces a concise, accurate listing of statistics that can be easily referenced.

INTRODUCTION

For a majority of the analyses that we routinely perform, we are primarily interested in comparing different groups in our patient populations. For certain outcomes, it is of interest to ascertain the size of any group differences while controlling for certain variables. More specifically, adjusted means, standard errors of the adjusted means, and p-values, which test the difference in the means, are of central interest. This paper describes a method that can be used to capture this key subset of statistics from PROC MIXED in a listing that can be quickly reviewed.

DESCRIPTION OF DATA

The example illustrated here is derived from a study of a Chronic Obstructive Pulmonary Disease (COPD) population that has been randomized to receive either standard care or the intervention. Twenty-four months of medical claims data were collected with the intention of determining any group differences in the change in charges associated with resource utilization. The first twelve months prior to study enrollment is considered a baseline assessment and the twelve months post-enrollment is referred to as the follow-up assessment. Charges were then totaled from each assessment period and subtracted. Other data were collected for each patient (i.e., marital status, smoking status, highest education achieved, and comorbidity index) and are used as covariates in our model. In addition, design strata (site and medical resource use classification) are also included, with site treated as a random effect.

PROGRAMMING STEPS

The following steps are needed to create the listing of pertinent statistics:

A. PROC DATASETS

PROC DATASETS is used to delete datasets from the work library that are no longer relevant. Its invocation is extremely important so that future datasets will be created anew, rather than overwritten. Since the majority of the SAS programming code is contained within a macro, different datasets having the same name will be repeatedly created. The potential for errors to unknowingly occur in this process increases if this PROC is missing from the code.

B. PROC MIXED

PROC MIXED creates the key statistics of interest and outputs them in datasets through ODS so that they can be manipulated in the DATA step. Since it is of interest to report the adjusted means and standard errors for each group, the LSMEANS syntax with the specification of DIFFS is incorporated into the PROC MIXED code. The MAKE LSMEANS syntax, which converts the table LSMEANS into a dataset, is required as well. To obtain p-values, the MAKE DIFFS statement is also needed. A second dataset results.

C. DATA step manipulations

The DATA step is used to merge datasets that were created in PROC MIXED. In addition, variables are created from the parameters that

©1999 Lewin-TAG, Inc.

were submitted to the macro. These variables are extremely important to describe the model that was used.

D. PROC APPEND

PROC APPEND is used to add the observations from one dataset to the end of another dataset. The resulting dataset contains the cumulative set statistics that have been output from the models that were invoked.

E. PROC DATASETS

PROC DATASETS is used to delete datasets from the work library. See A above.

F. Macro invocation

The macro invocation includes specification of the parameters that are to be used in the process.

G. PROC PRINT

PROC PRINT is used to produce a listing of key statistics.

SAS CODE

```
*** This step is only necessary if the program is
run more than once in the same SAS session;
proc datasets lib=work;
   delete temp05;
run;
quit;

%macro dowork( indat      = temp01,
               outcome    = ,
               subgroup   = ,
               what       = ,
               include    = );

   proc mixed noclprint data = &indat;
      where &subgroup = "&what";
      class group site marital smoker educate
            mruclass &include;
      model  &outcome =  group marital smoker
                         educate mruclass
                         &include;
      random int / subject=site;
      lsmeans group / diff om;
      make lsmeans
         out    = temp02(keep=group
                  _lsmean_ _se_);
      make diffs
         out    =  temp03(keep=_pt_);
run;
quit;

data temp04(keep=group outcome subgroup
                 adjmn stderror pvalue);
   length outcome subgroup $20.;
   set temp02;
   if _n_ = 1 then set temp03;

   outcome    = "&outcome";
   subgroup   = "&subgroup" || "-" || "&what";
   adjmn      = _lsmean_;
   stderror   = _se_;
   pvalue     = _pt_;

   format adjmn stderror 8.2 pvalue 8.4;

   label
      outcome    = 'Outcome'
      subroup    = 'Subgroup'
      adjmn      = 'Adjusted Mean'
      stderror   = 'Standard Error'
      pvalue     = 'P for Diff';
run;

proc append base = temp05 data = temp04;
run;

proc datasets lib = work;
   delete temp02-temp04;
run;
quit;

%mend dowork;

%dowork(  outcome    =erchg,
          subgroup   =all,
          what       =patients,
          include    =comorb);
%dowork(  outcome    =outchg,
          subgroup   =all,
          what       =patients,
          include    =comorb);
```

©1999 Lewin-TAG, Inc.

```
%dowork(  outcome   =totchg,
          subgroup  =all,
          what      =patients,
          include   =comorb);
%dowork(  outcome   =totchg,
          subgroup  =packs,
          what      =<= 40,
          include   =comorb);
%dowork(  outcome   =totchg,
          subgroup  =packs,
          what      =41-60,
          include   =comorb);
%dowork(  outcome   =totchg,
          subgroup  =packs,
          what      => 60,
          include   =comorb);
%dowork(  outcome   =totchg,
          subgroup  =comorb,
          what      =none,
          include   = );
%dowork(  outcome   =totchg,
          subgroup  =comorb,
          what      =one,
          include   = );
%dowork(  outcome   =totchg,
          subgroup  =comorb,
          what      => one,
          include   = );

proc print data = temp05 label;
   var group subgroup outcome adjmn stderror
       pvalue;
run;
```

(Note: This code is structured to accommodate subgroup analyses. To do group comparisons across the entire population, a constant character variable is required in the initial input dataset {temp01}. In this case, all = 'patients' for each record in the dataset. SAS then recognizes in the WHERE syntax of PROC MIXED that each patient is to be included in the analysis.)

RESULTS

obs	Group	Subgroup	Outcome	Adjusted Mean	Standard Error	P for Diff
1	CS	all-patients	erchg	54.93	33.17	0.8103
2	CSI	all-patients	erchg	65.73	29.51	0.8103
3	CS	all-patients	outchg	-3.61	21.59	0.7853
4	CSI	all-patients	outchg	4.37	19.20	0.7853
5	CS	all-patients	totchg	937.09	628.38	0.5447
6	CSI	all-patients	totchg	420.51	558.99	0.5447
7	CS	packs-<= 40	totchg	406.71	704.26	0.5432
8	CSI	packs-<= 40	totchg	-180.82	615.70	0.5432
9	CS	packs-41-60	totchg	-607.81	1506.76	0.2564
10	CSI	packs-41-60	totchg	1785.98	1303.05	0.2564
11	CS	packs-> 60	totchg	1600.90	1099.50	0.5187
12	CSI	packs-> 60	totchg	601.13	1056.93	0.5187
13	CS	comorb-none	totchg	471.26	363.84	0.5273
14	CSI	comorb-none	totchg	147.37	328.01	0.5273
15	CS	comorb-one	totchg	763.23	1020.81	0.5993
16	CSI	comorb-one	totchg	52.78	844.57	0.5993
17	CS	comorb-> one	totchg	1427.68	1022.01	0.7219
18	CSI	comorb-> one	totchg	918.21	978.97	0.7219

CONCLUSION

This method of obtaining and outputting key statistics from PROC MIXED has many benefits. Firstly, a concise listing is created. It is unnecessary to page through the numerous pages of output to locate those statistics of primary interest. Secondly, there is no need to print all of the PROC MIXED documentation. This saves a large amount of paper from having to be printed. And, thirdly, this method is flexible so that many different models can be run in one SAS session.

©1999 Lewin-TAG, Inc.

REFERENCES

SAS Institute Inc. *SAS/STAT Software, Changes and Enhancements through Release 6.11*. Cary, NC. SAS Institute Inc. 1990.

SAS Institute Inc. *SAS/STAT User's Guide, Version 6, Fourth Edition: Volume 2*. Cary, NC. SAS Institute Inc. 1990.

ACKNOWLEDGMENTS

SAS and SAS/STAT software are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

AUTHOR CONTACT

Stefanie Silva, Statistical Analyst
ssilva@lewin.com

Dave Miller, Senior Statistician
dmiller@lewin.com

Lewin-TAG, Inc.
490 2nd St., Suite 201
San Francisco, CA 94107
(415) 495-8966 (phone)
(415) 495-8669 (fax)

©1999 Lewin-TAG, Inc.