Using a SAS Macro to Create Unique Statistical Output

Sean W. Mulvenon, University of Arkansas, Fayetteville, AR
M. Austin Betz, Arizona State University, Tempe, AZ

Abstract

The SAS® Macro facility can be very useful in developing statistical output that is not available in any of the statistical procedures. The current paper attempts to provide an example of how the macro facility can be employed to provide specific or unique statistical output that is valuable to statisticians and behavioral science researchers. A new theoretical statistical result developed for selecting between a multivariate and univariate approach when analyzing repeated measures designs was selected for use in this paper. The mathematical intricacies and programming complexities involved in using this new statistical result may limit its actual use by both statisticians and researchers. Thus, two SAS macros have been developed to increase access and facilitate use of this new statistical result. The SAS procedures used are limited to SAS/IML and SAS/GLM and the working knowledge required to use these macros extends from the intermediate to the advanced user.

Introduction

Behavioral research often requires that subjects be tested on multiple occasions for assessing the effect of a treatment. One approach to assess the magnitude of treatment effects on subjects is to use a within-subjects or repeated measures design. A major advantage of within-subjects designs is that error variance attributable to individual differences is removed and the result is a statistical test that is more sensitive to group and subject differences while providing greater power. However, within-subjects designs can be analyzed using either a multivariate or univariate analysis of variance and this has produced a methodological debate as to which procedure should be employed.

Mulvenon & Betz (1997) developed a relationship between multivariate and univariate noncentrality parameters in repeated measures designs for generating prospective power estimates in the univariate procedure. Using a "bias ratio," upper and lower bounds for the univariate noncentrality parameter can be obtained (Boik, 1981). These bounds, incorporated with the necessary sample size to obtain a desired power level in the multivariate procedure, are used to compute upper and lower bound estimates of power in the univariate procedure. A comparison between the multivariate and univariate procedures is then completed with the conclusions providing guidelines, based on best and worst case scenarios, for selecting between the multivariate and univariate procedures.

The result from the univariate versus multivariate comparison provides the behavioral sciences researcher with more control when developing studies, allowing for *a priori* determination of sample size as a function of desired effect size, level of power, alpha, and the minimum correlation value of the data matrix. Although a decision on whether to use the univariate or multivariate is delayed until a variance-covariance matrix is computed, there is now a viable procedure that provides a format for this selection.

The paper by Mulvenon and Betz (1997) is very mathematically oriented and therefore the applied researcher may have difficulty in using the information and procedures. To address this issue and make the information more accessible to the applied behavioral science researcher, two SAS Macros were written and are provided in this paper to extend the accessibility of the information from the mathematical to the applied researcher.

SAS Macros

Using data for mixed model designs from Maxwell and Delaney (1990, p. 466) and the SAS manual for linear models (1991, p. 317)[1] some examples are provided to demonstrate how the applied researcher can use the various results discussed in Mulvenon and Betz (1997). The first example demonstrates a situation where a single sample repeated measured design is employed. The second example demonstrates a situation where a one-between and one-within repeated measured design is employed. Both examples demonstrate the use of macros to

---

[1] The data has been modified to create equal group sizes.

provide additional statistical information. Further, how the additional output can assist any researcher in *a priori* selection between the univariate or multivariate approach when developing a study.

Example 1

Single Sample Design. Suppose that a researcher is designing a study that will examine the effect of age on a child's cognitive abilities. The principal instrument that will be used is the McCarthy Scales of Children's Abilities (MSCA). Maxwell and Delaney (1990, p. 466) provided an example of data that might be obtained if 12 subjects were administered the MSCA on four occasions. Suppose further, that an effect size of $[(\Delta / \sigma) = 1.0]$, with power $[(1 - \beta) = .80]$ and $\alpha = .05$ are the parameter values for determining the necessary sample size for a desired level of power. First, a researcher could use the hypothetical data set in Maxwell and Delaney (1990, pg. 466) as a reference for estimating an expected minimum correlation. For example, the minimum correlation for the McCarthy Scales data is, $\rho_{min} = .4661$. Using the minimum correlation value of .4661 and by specifying the desired effect size, level of power, and the nominal level of alpha, the researcher can use the procedures in this study to determine the necessary sample size for the study being planned.

As a way of making this procedure more accessible to the behavioral sciences researcher, a SAS macro called SINGLE has been written that will generate retrospective and prospective power estimates. This macro, which can be accessed using and SAS job program, calculates the retrospective power for the referent data set and the prospective sample sizes for the desired parameter values (effect size, power level, and nominal alpha). An example program demonstrating this procedure can be seen in Appendix A, section I.

The normal SAS PROC/GLM output for the repeated measures analysis and the output produced by the Macro can be seen in Appendix A, section II. Retrospective results include sample size (N = 12), effect size (DELTA = 1.74), minimum correlation (RHO = .4661), Greenhouse-Geisser power (GG_PWR = .500), multivariate power (M_PWR =.153), departure from sphericity and EPSILON = .6095, univariate F-test (F-TEST = 3.03) and the multivariate test (M_HLT = 2.2401). Both the univariate F-test and the multivariate F-test are reported only as a frame of reference to compare with the standard SAS/GLM results.

Prospective results include the necessary sample size (TOTAL_N = 16), the desired effect size (PDELTA = 1.0), univariate lower and upper bound power estimaqtes (LU_PWR = .291, UU_PWR = .973), and multivariate power (PM_PWR = .806). The value TOTAL_N gives the total sample size required to attain the values defined in the command %SINGLE (PDELTA = 1.0, POWER = .80, ALPHA = .05).

The output reported in appendix A can be interpreted in a direct manner. The multivariate result using the value for the Hotelling-Lawley Trace (HLT) statistic is $F(3,9) = 2.2401$ $(p < 0.1528)$. The results of the univariate results F = 3.03 is a modified test using the Greenhouse-Geisser value of epsilon. The adjusted degrees of freedom produces an adjusted probability reported as $(p < .075)$. The power for the univariate or multivariate method is not very high, but the output does provide information that can aide the researcher in selecting the procedure that will provide greater power.

Examination of the prospective results can be used to determine the necessary sample size (TOTAL_N) for a future study with the desired effect size. More specifically, if a data set with characteristics similar to the model data set is collected, then using the results reported on in appendix A, 16 subjects would be necessary to attain the parameters specified in the command line. Further, the univariate procedure would be recommend as the appropriate statistical procedure. The lower bound estimate for the univariate case is only .291, but the upper bound estimate is .973. The expected effect size is larger than requested (based on the retrospective study), and thus, even the multivariate power is relatively large, the univariate procedure is likely to provide greater power.

Example 2

One-Between and One-Within Interaction Design. In this example the researcher is considering a study where each subject's heart rate is measured on seven occasions after performing a physical exercise (7 repeated trials). A treatment condition with three levels (control, weight increased, and exercise increased) is included. Thus, the number of groups is g = 3 and the number of trials is p = 7. Next, the researcher proceeds as in the first example with decisions on the minimum effect size, nominal alpha, and desired level of power. Once these values are selected, the researcher can use the macro SPLTPLOT[2] located in Appendix B, Section I.

The results from this analysis are reported in Appendix B, section II. The SAS GLM multivariate results for the TIME (WITHIN) and TIME*GROUP effects are F(6, 38) = 5.37 (p < .0004) and F(12, 74) = 1.33 (p < .22), respectively. The univariate results for the TIME and TIME*GROUP effects are F(6, 258) = 6.57 (p < .0010) and F(12, 258) = 3.01 (p < .0148), respectively. Retrospective power values are computed in the same fashion as those for example 1, with the addition of results for F_WITHIN, F_INTER, M_WITHIN, and M_INTER, representing the labels for the macro analogs of the statistical tests provided in the SAS output. Again, these values are only reported as a frame of reference to compare with those results generated using SAS GLM.

Prospective results include a variable for total sample size (TOTAL_N = 24), group size (GRP_SIZE = 8), desired effect size (PDELTA = 1.0), lower and upper bound prospective power estimates (LU_PWR = .05 and UU_PWR = 1.0) and a prospective multivariate power estimate (PM_PWR = .994). The results for the prospective power estimates, especially with PM_PWR = .994, indicates the multivariate procedure should be recommended for any future study with data that is expected to be similar to data for this example. The result supports a trend reported in Mulvenon and Betz (1997), that as the complexity of the design increases, the multivariate procedure will become consistently more powerful when compared to the univariate procedure. Thus, if a researcher is planning a study based on this output from this macro, he or she would want to use 24 subjects in each of 3 groups.

Conclusion

The purpose of this paper is to demonstrate how unique statistical output can be generated using SAS macros. Further, this paper also demonstrates how advanced or new statistical theory can be incorporated with standard statistical output produced by the various SAS Proc steps.

References

Boik, R. (1981). A priori tests in repeated measures designs: Effects of nonsphericity. Psychological Bulletin, 86, 1084-1089.

Littell, R., Freund, R., & Spector, P. (1991). SAS® System for Linear Models. SAS Institute: Cary, NC

Maxwell, S., & Delaney, H. (1990). Designing Experiments and Analzying Data: A Model Comparison Perspective. Wadsworth Publishing: Belmont, CA.

Mulvenon, S., & Betz, M.A. (1997). Determination of Prospective Sample Size Power Estimation in Repeated Measures Designs. Paper presented at the annual conference of the American Eductional Research Association, Chicago, IL.

SAS Institute Inc. (1990). SAS/IML® User's Guide (Version 6.06 ed.). Cary, NC: Author.

Appendix A

Section I: SINGLE Macro and SAS/GLM Program

SINGLE Macro

```
%MACRO SINGLE(PDELTA=, POWER=, ALPHA=);

PDELTA=&PDELTA;
POWER=&POWER;
ALPHA=&ALPHA;
N= NROW(X);
P= NCOL(X);
B= P-1;
C=1;
SUM= X(|+,|);
MEAN=SUM/N;
XPX= X`*X  - SUM`*SUM/N;
S= DIAG(1/SQRT(VECDIAG(XPX)));
CORRMAT= S*XPX*S;
COVMAT= XPX/(N-1);
U_I1=J(P,P-1,0);
DO I=1 TO P;
 DO J=1 TO B WHILE (J < I);
  U_I1(|I,J|)= -1;
  U_I1(|J,J|)=P - J;
 END;
END;
U_I1=U_I1/SHAPE(SQRT(U_I1[##,]),NROW(U_I1),NCOL(U_I1));
SIGMA_ST= U_I1`*COVMAT*U_I1;
M_SIGMA= SIGMA_ST*(N - 1);
EVAL1= EIGVAL(SIGMA_ST);
EPSILON= (SUM(EVAL1)*SUM(EVAL1))/((P - 1)*EVAL1`*EVAL1);
X_POP= VECDIAG(I(N));
THETA=C*MEAN*U_I1;
DELTA_ST=THETA`*INV(C*INV(X_POP`*X_POP)*C`)*THETA;
DELTA=SQRT(TRACE(DELTA_ST))/SQRT(TRACE(SIGMA_ST));
HLT= TRACE(DELTA_ST*INV(M_SIGMA));
F_TEST=TRACE(DELTA_ST)/TRACE(SIGMA_ST);
M_HLT= (HLT/(P-1))/(1/(N - P + 1));
RHO= MIN(CORRMAT);
```

---

[2] It should be noted that the macro SPLTPLOT is not in a generalized form. Thus, this program will work for only the 3 group case. A version that is a general form is under development.

```
/* INSERT VALUES FOR NECESSARY CONDITIONS IN PROSPECTIVE
   POWER ESTIMATION */

LB_EPS= 1/B;
     Q= P + 1;
    E1= 1/B;
    E2= (B - 1)/B;
    E3= 1/2*(1 + SQRT(1/B));
 E_MAX= 1/2*(1 + SQRT(B));
 E_MIN= 1/2*(1 - (B - 2)*SQRT(1/B));
IF E_MIN < .00 THEN E_MIN= .00;

/* GENERATE A RETROSPECTIVE MULTIVARIATE POWER ESTIMATE */

 M_NDF= (P - 1);             /* MULTIVARIATE NUMERATOR DF */
 M_DDF= (N - P + 1);         /* MULTIVARIATE DENOMINATOR DF */
 M_NCP= (M_DDf/M_NDF)*HLT;
 M_FCRIT= FINV(1-ALPHA, M_NDF, M_DDF);
 M_PWR= 1 - PROBF(M_FCRIT, M_NDF, M_DDF, M_NCP);

/* GENERATE A RETROSPECTIVE UNIVARIATE POWER ESTIMATE */

LB_EPS= 1/(P-1);
 GG_NDF= (P - 1)*EPSILON;     /* UNIVARIATE NUMERATOR DF */
 GG_DDF= (N - 1)*(P - 1)*EPSILON; /* UNIVARIATE DENOMINATOR DF */
 GG_NCP=B*EPSILON*TRACE(DELTA_ST)/TRACE(SIGMA_ST);
 GG_FCRIT= FINV(1-ALPHA, GG_NDF, GG_DDF);
 GG_PWR= 1 - PROBF(GG_FCRIT, GG_NDF, GG_DDF, GG_NCP);

     M_PWR= ROUND(M_PWR, .001);
     U_PWR= ROUND(GG_PWR, .001);

PRINT '*********************************************************************';
 PRINT '  ';
 PRINT 'POWER ANALYSIS RESULTS ';
 PRINT '  ';
 PRINT 'RESTROSPECTIVE: '   N DELTA RHO U_PWR M_PWR EPSILON
                                        F_TEST HLT M_HLT;

   /* GENERATE THE PROSPECTIVE MULTIVARIATE POWER  */

N3= P + 1;
DO N4= N3 TO 300 BY 1 UNTIL (PM_PWR > POWER);
 PM_NDF= (P - 1);              /* MULTIVARIATE NUMERATOR DF */
 PM_DDF= (N4 - P + 1);         /* MULTIVARIATE DENOMINATOR DF */
 PM_NCP= (N4*(PDELTA**2)) / (2*(1-RHO));
   IF PM_NCP > 400 THEN PM_NCP= 400;
 PM_FCRIT= FINV(1-ALPHA, PM_NDF, PM_DDF);
 PM_PWR= 1 - PROBF(PM_FCRIT, PM_NDF, PM_DDF, PM_NCP);
END;

LB_EPS= 1/(P-1);
EPS= E3;
TOTAL_N= N4;

   /* GENERATE THE UNIVARIATE UPPER BOUND ESTIMATE  */

    UU_NDF= (P-1)*EPS;
    UU_DDF= (TOTAL_N-1)*(P-1)*EPS;
    UU_NCP= PM_NCP*E_MAX;
     IF UU_NCP > 400 THEN UU_NCP= 400;
  UU_FCRIT= FINV(1-ALPHA, UU_NDF, UU_DDF);
    UU_PWR= 1 - PROBF(UU_FCRIT, UU_NDF, UU_DDF, UU_NCP);

   /* GENERATE THE UNIVARIATE LOWER BOUND ESTIMATE  */

    LU_NDF= (P-1)*EPS;
    LU_DDF= (TOTAL_N-1)*(P-1)*EPS;
    LU_NCP= PM_NCP*E_MIN;
     IF LU_NCP > 400 THEN LU_NCP= 400;
  LU_FCRIT= FINV(1-ALPHA, LU_NDF, LU_DDF);
    LU_PWR= 1 - PROBF(LU_FCRIT, LU_NDF, LU_DDF, LU_NCP);

    PM_PWR= ROUND(PM_PWR, .001);
    UU_PWR= ROUND(UU_PWR, .001);
    LU_PWR= ROUND(LU_PWR, .001);

IF (PM_PWR > POWER) & (UU_PWR >= PM_PWR) THEN DO;
 PRINT '  ';
 PRINT '    PROSPECTIVE:                  '  TOTAL_N PDELTA LU_PWR UU_PWR
     PM_PWR  '                              ';
 PRINT '  ';
PRINT '**********************************************************************';
END;
RUN;
%MEND SINGLE;
```

Section II: Statistial Output from SAS/GLM and Single Macro

SAS/GLM Program

```
DATA ONE; OPTIONS MAUTOSOURCE SASAUTOS= (`C:\MACROS' ) MRECALL
         LS=65 PS= 60 NODATE  PAGENO=1

INPUT SUBJECT AGE1 AGE2 AGE3 AGE4;
CARDSs;
 1 108  96 110 122
 2 103 117 127 133
 3  96 107 106 107
 4  84  85  92  99
 5 118 125 125 116
 6 110 107  96  91
 7 129 128 123 128
 8  90  84 101 113
 9  84 104 100  88
10  96 100 103 105
11 105 114 105 112
12 113 117 132 130
;

PROC GLM;
  MODEL AGE1-AGE4 = / NOUNI;
  REPEATED TIME;
TITLE2 'SINGLE SAMPLE REPEATED MEASURES';
RUN;

PROC IML;
USE WORK.ONE;
READ ALL VAR {AGE1 AGE2 AGE3 AGE4}INTO X;

%SINGLE (PDELTA= 1.0, POWER= .80, ALPHA= .05);
```

Statistical Output

```
                         The SAS System
                 SINGLE SAMPLE REPEATED MEASURES

                 General Linear Models Procedure
             Repeated Measures Analysis of Variance
               Repeated Measures Level Information

    Dependent Variable       AGE1     AGE2     AGE3     AGE4

         Level of TIME          1        2        3        4

        Manova Test Criteria and Exact F Statistics for
                the Hypothesis of no TIME Effect
   H = Type III SS&CP Matrix for TIME   E = Error SS&CP Matrix

                      S=1    M=0.5    N=3.5

Statistic                Value        F    Num DF  Den DF  Pr > F
Wilks' Lambda           0.572508   2.2401       3       9  0.1528
Pillai's Trace          0.427492   2.2401       3       9  0.1528
Hotelling-Lawley Trace  0.746699   2.2401       3       9  0.1528
Roy's Greatest Root     0.746699   2.2401       3       9  0.1528

                         The SAS System                         3
                 SINGLE SAMPLE REPEATED MEASURES

                 General Linear Models Procedure
             Repeated Measures Analysis of Variance
    Univariate Tests of Hypotheses for Within Subject Effects

Source: TIME
                                                    Adj  Pr > F
     DF Type III SS Mean Square  F Value  Pr > F   G - G   H - F
      3   552.00000   184.00000     3.03  0.0432  0.0748  0.0635

Source: Error(TIME)

     DF Type III SS Mean Square
     33  2006.00000    60.78788

               Greenhouse-Geisser Epsilon = 0.6095
                      Huynh-Feldt Epsilon = 0.7249





                         The SAS System                         4
                 SINGLE SAMPLE REPEATED MEASURES

                 ***********************************************
                     POWER ANALYSIS RESULTS
```

```
                      N     DELTA       RHO       U_PWR
RESTROSPECTIVE:       12 1.7398044  0.4660875      0.5

                   M_PWR   EPSILON  F_TEST      HLT      M_HLT
                   0.153 0.6095445 3.0269192 0.7466993 2.2400979

                    TOTAL_N    PDELTA    LU_PWR
   PROSPECTIVE:       16         1       0.291

                     UU_PWR    PM_PWR
                     0.973     0.806

                ********************************************************
```

Appendix B

<u>**Section I**</u>: SPLTPLOT Macro and SAS/GLM Program

SPLTPLOT Macro

```
%MACRO SPLTPLOT(PDELTA=, POWER=, ALPHA=, GROUPS=);

GROUPS=&GROUPS;
PDELTA=&PDELTA;
POWER=&POWER;
ALPHA=&ALPHA;
N= NROW(X);
P= NCOL(X);
X= X(|1:N, 1:P|);
B= P-1;
R=&GROUPS;
N1= N/R;
A= R-1;
SUM=X(|+,|);
XPX=X`*X - SUM`*SUM/N;
S= DIAG(1/SQRT(VECDIAG(XPX)));
CORRMAT= CORR(X);
COVMAT1= S*CORRMAT*S;
COVMAT= XPX/(N-1);
CMATRIX1= VECDIAG(I(R));
 CMATRIX2=J(R, R-1, 0);
 DO G= 1 TO R;
  DO H= 1 TO A WHILE (H < G);
   CMATRIX2(|G,H|)= -1;
   CMATRIX2(|H,H|)= R - H;
  END;
 END;
CMATRIX1=T(CMATRIX1);
CMATRIX2=T(CMATRIX2);
U_I1=J(P,P-1,0);
DO I=1 TO P;
 DO J=1 TO B WHILE (J < I);
  U_I1(|I,J|)= -1;
  U_I1(|J,J|)=P - J;
 END;
END;
U_I1=U_I1/SHAPE(SQRT(U_I1[##,]),NROW(U_I1),NCOL(U_I1));
X1= X(|1:N1, 1:P |);
X2= X(|N1+1:2*N1, 1:P|);
X3= X(|2*N1+1:3*N1, 1:P|);
SUM1= X1(|+,|);
SUM2= X2(|+,|);
SUM3= X3(|+,|);
MEAN1=SUM1/N1;
MEAN2=SUM2/N1;
MEAN3=SUM3/N1;
MEAN=MEAN1//MEAN2//MEAN3;
G=NROW(MEAN);
XPX1= X1`*X1  - SUM1`*SUM1/N1;
XPX2= X2`*X2  - SUM2`*SUM2/N1;
XPX3= X3`*X3  - SUM3`*SUM3/N1;
S1= XPX1/(N1-1);
S2= XPX2/(N1-1);
S3= XPX3/(N1-1);
S1_ST= U_I1`*S1*U_I1;
S2_ST= U_I1`*S2*U_I1;
S3_ST= U_I1`*S3*U_I1;
SIGMA= U_I1`*COVMAT*U_I1;
SIGMA_ST= ((N1-1)*(S1_ST+S2_ST+S3_ST))/(N-R);
M_SIGMA= SIGMA_ST*(N - R);
EVAL1= EIGVAL(SIGMA_ST);
EPSILON= (SUM(EVAL1)*SUM(EVAL1))/((P - 1)*EVAL1`*EVAL1);
X_POP= I(R)@SHAPE(1,N1,1);
THETA1= CMATRIX1*MEAN*U_I1;
DELTA1= THETA1`*INV(CMATRIX1*INV(X_POP`*X_POP)*
        CMATRIX1`)*THETA1;
THETA=CMATRIX2*MEAN*U_I1;
DELTA_ST=THETA`*INV(CMATRIX2*INV(X_POP`*X_POP)*
         CMATRIX2`)*THETA;
DELTA=SQRT(TRACE(DELTA_ST))/SQRT(TRACE(SIGMA_ST));
HLT= TRACE(DELTA_ST*INV(M_SIGMA));
HLT1= TRACE(DELTA1*INV(M_SIGMA));
S= A >< B;
M_WITHIN= (HLT1/(P-1))/(1/(N -P -R + 2));
M_HLT= ((HLT/S)/(A*B))/(1/(S*((N-R) - B - 1) + 2));
F_WITHIN=TRACE(DELTA1)/TRACE(SIGMA_ST);
F_INTER=(TRACE(DELTA_ST)/((P-1)*(G-1)))/(TRACE(SIGMA_ST)/(P-1));
RHO= MIN(CORRMAT);

/* INSERT VALUES FOR NECESSARY CONDITIONS IN PROSPECTIVE
   POWER ESTIMATION */

LB_EPS= 1/B;
     Q= GROUPS*(P + 1);
    E1= 1/B;
    E2= (B - 1)/B;
    E3= 1/2*(1 + SQRT(1/B));
   EPS= E3;
 E_MAX= EPS + SQRT((B - 1)*(1 - EPS)*EPS);
 E_MIN= EPS - SQRT((B - 1)*(1 - EPS)*EPS);
 IF E_MIN < .00 THEN E_MIN= .00;

/* GENERATE A RETROSPECTIVE MULTIVARIATE POWER ESTIMATE */

 M_NDF= (P - 1)*(G - 1);  /* MULTIVARIATE NUMERATOR DF */
 M_DDF= S*((N - G) - (P - 1) - 1) + 2; /*MULTIVARIATE DENOM DF */
 M_NCP= (M_DDF/S)*HLT;
 M_FCRIT= FINV(1-ALPHA, M_NDF, M_DDF);
 M_PWR= 1 - PROBF(M_FCRIT, M_NDF, M_DDF, M_NCP);

/* GENERATE A RETROSPECTIVE UNIVARIATE POWER ESTIMATE */

 GG_NDF= (P - 1)*(G - 1)*EPSILON;     /* UNIVARIATE NUMERATOR DF */
 GG_DDF= (P - 1)*(N - G)*EPSILON;  /* UNIVARIATE DENOMINATOR DF */
 GG_NCP=B*EPSILON*TRACE(DELTA_ST)/TRACE(SIGMA_ST);
 GG_FCRIT= FINV(1-ALPHA, GG_NDF, GG_DDF);
 GG_PWR= 1 - PROBF(GG_FCRIT, GG_NDF, GG_DDF, GG_NCP);

     M_PWR= ROUND(M_PWR, .001);
     U_PWR= ROUND(GG_PWR, .001);

 PRINT '*********************************';
 PRINT '  ';
 PRINT 'POWER ANALYSIS RESULTS ';
 PRINT '  ';
 PRINT 'RETROSPECTIVE: ' N DELTA RHO U_PWR M_PWR EPSILON
                                 F_WITHIN M_WITHIN F_INTER M_HLT;

   /* CALCULATE A PROSPECTIVE MULTIVARIATE POWER ESTIMATE */

DO N3= Q TO 1000 BY 3 UNTIL (PM_PWR > POWER);
     PM_NDF= A*B;
     PM_DDF= S*(N3 - GROUPS - B - 2) + 2;
     PM_NCP= (N3*(PDELTA**2))/(2*(1 - RHO));
      IF PM_NCP > 400 THEN PM_NCP= 400;
   PM_FCRIT= FINV(1-ALPHA, PM_NDF, PM_DDF);
     PM_PWR= 1 - PROBF(PM_FCRIT, PM_NDF, PM_DDF, PM_NCP);
END;

TOTAL_N=N3;
GRP_SIZE= TOTAL_N/GROUPS;

   /* GENERATE THE UNIVARIATE UPPER BOUND ESTIMATE  */

    UU_NDF= A*B*EPS;
    UU_DDF= B*(N3-GROUPS)*EPS;
    UU_NCP= PM_NCP*E_MAX;
     IF UU_NCP > 400 THEN UU_NCP= 400;
  UU_FCRIT= FINV(1-ALPHA, UU_NDF, UU_DDF);
    UU_PWR= 1 - PROBF(UU_FCRIT, UU_NDF, UU_DDF, UU_NCP);

   /* GENERATE THE UNIVARIATE LOWER BOUND ESTIMATE  */

    LU_NDF= A*B*EPS;
    LU_DDF= B*(N3-GROUPS)*EPS;
    LU_NCP= PM_NCP*E_MIN;
     IF LU_NCP > 400 THEN LU_NCP= 400;
  LU_FCRIT= FINV(1-ALPHA, LU_NDF, LU_DDF);
    LU_PWR= 1 - PROBF(LU_FCRIT, LU_NDF, LU_DDF, LU_NCP);

    PM_PWR= ROUND(PM_PWR, .001);
    UU_PWR= ROUND(UU_PWR, .001);
    LU_PWR= ROUND(LU_PWR, .001);

 IF (PM_PWR > POWER) & (UU_PWR >= PM_PWR) THEN DO;
 PRINT '  ';
 PRINT 'PROSPECTIVE:        ' TOTAL_N GRP_SIZE PDELTA LU_PWR
```

```
    UU_PWR PM_PWR      '            ';
 PRINT ' ';
 PRINT '*******************************';
 END;
RUN;
%MEND SPLTPLOT;
```

Section II: Statistial Output from SAS/GLM and SPLTPLOT Macro

SAS/GLM Program

```
DATA ONE; OPTIONS MAUTOSOURCE SASAUTOS= ('C:\MACROS' ) MRECALL
          LS= 65 PS=60 NODATE PAGENO=1;
iINFILE "C:\QUANT\SPLTPLOT.DAT";
INPUT GROUP S1 S2 S3 S4 S5 S6 S7;

PROC GLM;
  CLASS GROUP;
  MODEL S1-S7 = GROUP / NOUNI;
  REPEATED TIME;
RUN;

PROC IML;
USE WORK.ONE;
READ ALL VAR {S1 S2 S3 S4 S5 S6 S7}INTO X;

%SPLTPLOT (GROUPS= 3, PDELTA= 1.0, POWER= .80, ALPHA= .05);
```

SAS OUTPUT

```
                         The SAS System                       2
                 SINGLE SAMPLE REPEATED MEASURES

                 General Linear Models Procedure
             Repeated Measures Analysis of Variance
               Repeated Measures Level Information

Dependent Variable         S1       S2       S3       S4

     Level of TIME          1        2        3        4

Dependent Variable         S5       S6       S7

     Level of TIME          5        6        7

        Manova Test Criteria and Exact F Statistics for
                the Hypothesis of no TIME Effect
   H = Type III SS&CP Matrix for TIME   E = Error SS&CP Matrix

                       S=1    M=2    N=19

Statistic                Value        F    Num DF  Den DF  Pr > F
Wilks' Lambda           0.541132   5.6532       6      40  0.0003
Pillai's Trace          0.458868   5.6532       6      40  0.0003
Hotelling-Lawley Trace  0.847978   5.6532       6      40  0.0003
Roy's Greatest Root     0.847978   5.6532       6      40  0.0003

         Manova Test Criteria and F Approximations for
             the Hypothesis of no TIME*GROUP Effect
H = Type III SS&CP Matrix for TIME*GROUP   E = Error SS&CP Matrix

                      S=2    M=1.5    N=19

Statistic                Value        F    Num DF  Den DF  Pr > F

Wilks' Lambda           0.710363   1.2432      12      80  0.2693
Pillai's Trace          0.308405   1.2458      12      82  0.2671
Hotelling-Lawley Trace   0.38131   1.2393      12      78  0.2724
Roy's Greatest Root     0.290299   1.9837       6      41  0.0902

  NOTE: F Statistic for Roy's Greatest Root is an upper bound.
          NOTE: F Statistic for Wilks' Lambda is exact.
```

```
                         The SAS System                       4
                 SINGLE SAMPLE REPEATED MEASURES

                 General Linear Models Procedure
             Repeated Measures Analysis of Variance
```

```
    Univariate Tests of Hypotheses for Within Subject Effects

Source: TIME
                                                    Adj  Pr > F
     DF Type III SS Mean Square  F Value  Pr > F   G - G   H - F
      6   57.202381    9.533730     7.62  0.0001  0.0003  0.0002

Source: TIME*GROUP
                                                    Adj  Pr > F
     DF Type III SS Mean Square  F Value  Pr > F   G - G   H - F
     12   44.458333    3.704861     2.96  0.0007  0.0153  0.0119

Source: Error(TIME)

     DF Type III SS Mean Square
    270  337.767857    1.250992

               Greenhouse-Geisser Epsilon = 0.4132
                      Huynh-Feldt Epsilon = 0.4586

                         The SAS System                       5
                 SINGLE SAMPLE REPEATED MEASURES

                 *******************************
                     POWER ANALYSIS RESULTS

                        N     DELTA       RHO     U_PWR     M_PWR
RETROSPECTIVE:         48 2.4337372 0.7701305     0.839     0.649

                 EPSILON  F_WITHIN  M_WITHIN  F_INTER     M_HLT
                0.4132121 7.6209358 5.6531844 2.9615385 1.2392581

                   TOTAL_N  GRP_SIZE    PDELTA    LU_PWR
 PROSPECTIVE:         24         8         1      0.05

                    UU_PWR    PM_PWR
                      1        0.994

                 *******************************
```

SAS and SAS/IML are registered trademarks or trademarks of SAS institute Inc. In the USA and other countries. ® indicates USA registration.

Sean W. Mulvenon, Ph.D.
241 Graduate Education Building
University of Arkansas
Fayetteville, AR 72701
E-mail: seanm@comp.uark.edu