A FLEXIBLE MEAN-STANDARD ERROR PLOT: MACRO PROGRAM

WILL A. SULLIVAN, MERRILL DOW RESEARCH INSTITUTE
LIANN YUH, MERRILL DOW RESEARCH INSTITUTE

Abstract

A flexible mean-standard error plot macro program was developed using the HILOT
and HILOC options in PROC GPLOT, Version 5. Using these options, the means of two
treatment groups can be plotted with standard
error bars so that the bars do not interfere
with each other. This is true even if
error bars for each treatment at any
specifying the direction of the bars so that one bar does
not interfere with the other.

In SAS, one can use the I-STD option
in PROC GPLOT to obtain two-sided
mean-standard error plots, but this can
create a cluttered look when the bars
overlap; A procedure using SAS macros is
developed in Section 2. The macros
automatically generate one-sided or two-sided
error bars for each treatment at any
specified time points. Section 3 illustrates
the application of the macros with two
biological data sets. Concluding remarks as
well as possible modifications for more-than
two treatment cases are contained in Section
4.

[Sec 1] Introduction

Mean-standard error plots are demanded
for many types of data analyses. In
particular, the mean-standard error plot for
two treatments at some specified time points
or dose levels. Most of the existing graphic
packages can produce mean plots with
one-sided or two-sided standard error bars;
yet few have the ability to choose the
direction of the bars so that one bar does
not interfere with the other.

In SAS, one can use the I-STD option
in PROC GPLOT to obtain two-sided
mean-standard error plots, but this can
create a cluttered look when the bars
overlap; A procedure using SAS macros is
developed in Section 2. The macros
automatically generate one-sided or two-sided
error bars for each treatment at any
specified time points. Section 3 illustrates
the application of the macros with two
biological data sets. Concluding remarks as
well as possible modifications for more-than
two treatment cases are contained in Section
4.

[Sec 2] Macro Development

Computer: IBM 4381 SAS Version 5

The SAS macros were developed to
automatically create a mean plot with
standard error bars for two treatment or
response groups. The standard error bars are
plotted in one or both direction(s) depending
upon the magnitude of the mean values (+
standard errors). To create this type plot,
one accesses a partitioned data set (PDS)
with a JCL statement. The PDS contains the
macros which manufacture the plot and will be
thought of by SAS as the current macro
call library.

e.g. //FD DD DSN=DATA SET NAME, DISP=SHR
In addition to the usual OPTIONS
parameters, the OPTIONS statement must be
included and contain the parameters
MATDOSOURCE and SASTATUS=.
SASTATUS= specifies the file reference that the macro
facility uses as the current autocall library
and MATDOSOURCE directs the macro facility to
look in the autocall library for a member
with the same name as the macro that is
accessed.

Some initialization of the data is
required. The variable TRT TPLT must be added
to the data one wants plotted.

Treatment or Response (1): TRT TPLT='A';
Treatment or Response (2): TRT TPLT='B';

After the data has the variable
"TRT PLT" as a member then nine "LET"'s have
to be initialized. They are:

%LET TRT=;
%LET FILENAME=;
%LET YVAR=;
%LET YAXIS=;
%LET XYLABEL=;
%LET TITLE=;
%LET XLABEL=;
%LET XAXIS=;
%LET SLETS=;

Next add the title(s) for the plot in
SAS graphics notation and then call the macro
driver XPLTDRIVR and the mean plot will
automatically become a member in the graphics
catalog specified in

XLET FILENAME=;

For an example, here are the SAS statements to produce the attached plot.
Remember, the variable "TRT TPLT" has been
added to the data observations.

%LET TRT=; "A";
%LET TITLE= "A FLEXIBLE MEAN-STANDARD ERROR PLOT: MACRO PROGRAM";

... (some additional SAS code for initializing variables)

%MACRO XPLTDRIVR;
%* This is the driver macro to call

%MACRO THEPLOT;
1. DIVIDE
2. FINDPTS
3. TOME<GE
4. THEPLOT

After macro THEPLOT, a plot will be
produced with mean points plotted with
STDERR bars that do not overlap;

*DIVIDE
*FINDPTS
*TOME<GE
*THEPLOT

%MEND XPLTDRIVR;

503
The raw data set is passed by the driver macro to another macro DIVIDE which creates a data set using PROC MEANS (TODIVIDE) that contains the mean and standard errors for each treatment at each specified time point. TODIVIDE is divided into two separate data sets, one for each treatment. The response variables are renamed for merging purposes and keeping the original mean file intact.

**MACRO FINDPTS:**

1. * FIND THE MEAN AND STDERR OF THE DATA TO BE PLOTTED THEN OUTPUT THE DATA TO BE DIVIDED INTO TWO DATA SETS FOR MERGING AND COMPARISON IN ANOTHER MACRO FINDPTS;
2. DATA TPLT: SET DATAPLOT;
   PROC SORT DATA=TPLT: BY &XVAR:
   PROC MEANS DATA=TPLT: BY &XVAR:
   VAR STDERR;
   OUTPUT OUT=FINDPTS;
3. DATA TPLT: (DROP<CONC STDERR CONC2 STDERR2)
   PROC SORT DATA=TODIVIDE: BY &XVAR:
   IF TRT_TPLT=B THEN DO;
   CONC1=CONC2;
   STDERR1=STDERR2;
   OUTPUT TPLT_B;
   END;
   IF TRT_TPLT=A THEN DO;
   CONC1<CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_A;
   END;
   IF (CONC1-STDERR1) > (CONC2-STDERR2) THEN DO;
   CONC1<CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_B;
   END;
   IF (CONC1-STDERR1) < (CONC2-STDERR2) THEN DO;
   CONC1>CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_A;
   END;
   IF (CONC1-CONC2) > (STDERR1+STDERR2) THEN DO;
   OUTPUT TPLT_B;
   END;
   IF (CONC1-CONC2) < (STDERR1+STDERR2) THEN DO;
   OUTPUT TPLT_A;
   END;

**MACRO DIVIDE:**

1. * THIS MACRO WILL OUTPUT THE POINTS TO BE PLOTTED INTO 4 DATA SETS: TPLT_A1 AND TPLT_B1 WILL CONTAIN POINTS TO BE PLOTTED IN THE HILOT OPTION BECAUSE OF CROSSING STDERR BARS. TPLT_A2 AND TPLT_B2 WILL CONTAIN THE POINTS TO BE PLOTTED IN THE HILOE OPTION BECAUSE THEIR STDERR BARS DO NOT CROSS;
2. DATA TPLT_A: (DROP<CONC STDERR CONC2 STDERR2)
   TPLT_B: (DROP<CONC STDERR CONC1 STDERR1);
   SET FINDPTS;
3. IF TRT_TPLT=B THEN DO;
   CONC1<CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_B;
   END;
   IF TRT_TPLT=A THEN DO;
   CONC1<CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_A;
   END;
   IF (CONC1-STDERR1) > (CONC2-STDERR2) THEN DO;
   CONC1<CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_B;
   IF (CONC1-STDERR1) < (CONC2-STDERR2) THEN DO;
   CONC1>CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_A;
   END;

**MACRO FINDPTS:**

1. * THIS MACRO WILL OUTPUT THE POINTS TO BE PLOTTED INTO 4 DATA SETS: TPLT_A1 AND TPLT_B1 WILL CONTAIN POINTS TO BE PLOTTED IN THE HILOT OPTION BECAUSE OF CROSSING STDERR BARS. TPLT_A2 AND TPLT_B2 WILL CONTAIN THE POINTS TO BE PLOTTED IN THE HILOE OPTION BECAUSE THEIR STDERR BARS DO NOT CROSS;
2. DATA TPLT_A: TPLT_B:
   TPLT_A2: TPLT_B2: MERGE TPLT_A TPLT_B: BY &XVAR;
   PROC SORT DATA=TPLT_A2: BY &XVAR;
   IF previous points to help check or the direction of the.stderr bars;
   RETAIN P CONC1 P CONC2;
   * INITIATE THE PREVIOUS POINTS;
   IF _N_=1 THEN DO;
   CONC1=CONC2;
  RootElement="CONC1=CONC2:"
   OUTPUT TPLT_A2;
   END;
   ELSE IF P CONC1 LT P CONC2 THEN DO;
   OUTPUT TPLT_A2;
   CONC1=CONC2;
   STDERR1=STDERR2;
   OUTPUT TPLT_B2;
   END;
   ELSE IF P CONC1 GT P CONC2 THEN DO;
   OUTPUT TPLT_B2;
   CONC1=CONC2;
   STDERR1=STDERR2;
   OUTPUT TPLT_A2;
   END;
   IF the first assigned treatment variable is larger than the second then check to see if the stderr bars overlap. If the bars do not overlap then output the points to their respective data sets otherwise output the data points to the data sets which will plot stderr bars in only one direction;
2. IF CONC1 GT CONC2 THEN DO;
   IF CONC1=STDERR1 THEN DO;
   OUTPUT TPLT_B2;
   CONC1=CONC2;
   STDERR1=STDERR2;
   OUTPUT TPLT_A2;
   END;
   ELSE IF CONC1 LT CONC2 THEN DO;
   IF CONC1=STDERR1 THEN DO;
   OUTPUT TPLT_A2;
   CONC1=CONC2;
   STDERR1=STDERR2;
   OUTPUT TPLT_B2;
   ELSE DO;
   OUTPUT TPLT_A2;
   CONC1<CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_B2;
   END;
   ELSE DO;
   OUTPUT TPLT_B2;
   CONC1>CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_A2;
   END;

**MACRO FINDPTS:**

1. * DO THE SAME FOR AS ABOVE IF THE SECOND ASSIGNED TREATMENT VARIABLE IS LARGER THAN THE FIRST;
2. ELSE IF CONC1 LT CONC2 THEN DO;
   IF CONC1=STDERR1 THEN DO;
   OUTPUT TPLT_A2;
   CONC1=CONC2;
   STDERR1=STDERR2;
   OUTPUT TPLT_B2;
   ELSE DO;
   OUTPUT TPLT_B2;
   CONC1<CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_A2;
   END;
   ELSE DO;
   OUTPUT TPLT_B2;
   CONC1<CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_A2;
   END;

**MACRO FINDPTS:**

1. * IF THE TREATMENT VARIABLES ARE EQUAL THEN CHECK THE PREVIOUS POINTS TO SEE IN WHAT DIRECTION AND MAGNITUDE THE STDERR BARS WILL APPEAR. IF THE PREVIOUS POINTS WERE ALSO EQUAL THEN OUTPUT THE TIGHTEST INTERVAL (IF SMALLEST STDERR OF THE TWO VARIABLES) ;
2. ELSE DO;
   IF P CONC1 LT P CONC2 THEN DO;
   OUTPUT TPLT_A2;
   CONC1<CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_B2;
   ELSE DO;
   OUTPUT TPLT_B2;
   CONC1<CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_A2;
   END;
   ELSE DO;
   OUTPUT TPLT_A2;
   CONC1>CONC2;
   STDERR1>STDERR2;
   OUTPUT TPLT_B2;
   END;
   ELSE DO;
   OUTPUT TPLT_B2;
   CONC1>CONC2;
   STDERR1<STDERR2;
   OUTPUT TPLT_A2;
   END;

504
%MACRO FINDPTS:
%* RESET THE PREVIOUS POINTS;
P CONC=P CONC;
P CONC=CONC=CONC2;
%MEND FINDPTS;

--%TOMERGE is the next macro called by the driver macro and merges all of the data into one data set. New response, independent, and treatment variables are also created in each data set to achieve the final plotting purpose.

DATA SET 1 Means Treatment 'A'
DATA SET 2 Means Treatment 'B'
DATA SET 3 Observations Treatment 'A'
DATA SET 4 Observations Treatment 'B'
DATA SET 5 Observations Treatment 'A'
DATA SET 6 Observations Treatment 'B'

%MACRO TOMERGE:
%* THIS MACRO ASSIGN TREATMENT GROUPS FOR PLOTTING ALL THE DATA AND STDERR BARS. ONLY THREE VARIABLES WILL BE KEPT: TRT .. TREATMENT ASSIGNED FOR SORTING AND SYMBOL STATEMENTS CONC .. POINTS TO BE PLOTTED HOUR,. XAXIS USUALLY HOURS POST DOSING THEN ALL THE DATA WILL BE MERGED INTO ONE DATA SET FOR PLOTTING;

DATA TRT A; SET TRT A;
KEEP TRT CONC HOUR;
CONC=CONC;
PROC SORT; BY TRT HOUR;

DATA TRT B; SET TRT B;
KEEP TRT CONC HOUR;
CONC=CONC;
PROC SORT; BY TRT HOUR;

DATA TRT A2; SET TRT A2;
KEEP TRT CONC HOUR;
CONC=CONC;
PROC SORT; BY TRT HOUR;

DATA TRT B2; SET TRT B2;
KEEP TRT CONC HOUR;
CONC=CONC;
PROC SORT; BY TRT HOUR;

DATA ALLPTS; SET TRT_A TRT_B TRT_A2 TRT_B2;

%MEND TOMERGE;

THEPLOT contains all the SAS code from which the plot is created and placed in a SAS graphics catalog. The macro uses LETs to label the X-axis and Y-axis. They are also used with a FOOTNOTE statement to make a legend at the bottom of the plot.

The mean data are plotted first using the first two SYMBOL statements and a line is used to join the symbols. Next, the one-sided bars are drawn using the HILOC option in PROC GLOT. Finally, the two-sided bars are generated using the HILOT option in PROC GLOT.

%MACRO THEPLOT;
%* THIS MACRO WILL PLOT THE INFORMATION INTO THE OWNER'S GRAPHICS CATALOG. THIS WILL BE ACCOMPLISHED WITH THE USE OF THE PERCENT LETS THAT THE OWNER OF THE PLOT WILL PROVIDE:
GOPTIONS COLORS=(BLACK);
PROC GPLOT DATA=ALLPTS GOUT='&FILENAME';
AXIS1 ORDER='XAXIS' VALUE='(H=1 F=SIMPLEX) LABEL=(H=90 R=90 H .. DUPLEX &XLABEL) MINOR=NONE;
AXIS2 ORDER='XAXIS' VALUE='(H=1 F=SIMPLEX) LABEL=(A=90 R=0 H .. DUPLEX &YLABEL) MINOR=NONE;
PLOT CONC*HOUR-TRT/ HAXIS=AXIS1 VAXIS=AXIS2 NOLEGEND;
SYMBOL1 H=1 F=SPECIAL V=H I=J;
SYMBOL2 H=1.5 F=SPECIAL V=J I=J;
SYMBOL3 H=1 I=HILLO V=H;
SYMBOL4 H=1 I=HILLO V=H;
SYMBOL5 H=1 I=HILLO V=H;
SYMBOL6 H=1 I=HILLO V=H;
FOOTNOTE1 J=C H .. 1 F=SPECIAL 'J' F=DUPLEX &TRT1 F=DUPLEX &TRT2;
%MEND THEPLOT;

[Sec. 3] Examples
Two biological examples are used to demonstrate our macro program. The first example is obtained from a bioequivalence study. There is one plasma concentration-time profile for each formulation (treatment) and the graph is depicted in Figure 1.
The next example shows one experimental compound with two dose levels (treatments). The response variable is heart rate change from baseline and the X-axis variable is a time unit post dosing. The result is shown in Figure 2.

Both graphs clearly avoid the overlapping standard error bars even if the treatment means cross at any point in the plots. Furthermore, the titles, footnotes, symbols, labels and colors can easily be modified to suit special needs.

Possible modifications to generate this type plot for more than two treatments has also been studied. However, it is extremely difficult to define the criteria when error bars overlap, treatment means are equal or crossed at the same time point. We are still looking into this as well as refining our procedure.

References