TIPS TO ENHANCE YOUR SAS® STATISTICAL GRAPHICS OUTPUT

Shi-Tao Yeh, GlaxoSmithKline, King of Prussia, PA

ABSTRACT

SAS/GRAPH® software procedures produce several types of graphics output from a procedure. The procedures require a few statements to produce statistical graphics output.

Statistical graphics illustrate trends and test results. Many things can be done to improve your SAS® graphics output.

This paper provides tips to enhance your SAS statistical graphics output with sample code and illustrations.

The SAS products used in this paper are SAS BASE® and SAS/GRAPH on UNIX platform.

INTRODUCTION

The ancient Chinese proverb "a picture is worth a thousand words" describes the power of a picture. Graphics are pictures and are often a good means of communicating information. Statistical graphics represent statistical information and illustrate or compare trends. Good pictures offer clearly defined statistical results and elegantly represented information. It also offers sufficient detail for optimum clarity and eliminates confusion.

The procedures require a few statements to produce output. Most SAS/GRAPH procedures produce the following graphics output formats:

* graphics displayed on a graphics device,
* pictures stored in a graphics catalog,
* device commands sent to a sequential graphics stream file (GSF).

The SAS graph sample code in this paper produces a GSF file.

The majority of statistical graphs present in a coordinate system and use various elements such as points, lines, curves, bars, labels, titles, footnotes and other graphics primitives. SAS statements can control these elements.

This paper provides tips to enhance your SAS graphics output with sample code and illustrations.

DATA FILE AND SAMPLE OUTPUT

The data used for the illustration purposes is the plasma concentration time curve from D.W. Borne (1995). The data listing for the plasma concentration is shown as follows:

<table>
<thead>
<tr>
<th>TIME (HRS)</th>
<th>CONCENTRATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1 Time vs. Plasma Concentration Data

The function form for these types of curves is specified as:

\[ Y = p_1 \cdot (1 - \exp(-k \cdot (X - c))) \quad \ldots (1) \]

where:

\[ Y = \text{the concentration at time } X, \]
\[ X = \text{time in specified unit,} \]
\[ p_1 = \text{asymptote of the curve,} \]
\[ k = \text{rate constant,} \]
\[ c = \text{lag time in time unit}. \]

The nonlinear regression is run estimating for parameters \( p_1, k, \) and \( c. \) The estimation results and predicted values are shown in the Tables 2 and 3.

<table>
<thead>
<tr>
<th>TIME</th>
<th>CONCENTRATION</th>
<th>PREDICTED VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100.0</td>
<td>100.152</td>
</tr>
<tr>
<td>1</td>
<td>71.0</td>
<td>70.722</td>
</tr>
<tr>
<td>2</td>
<td>50.0</td>
<td>49.929</td>
</tr>
<tr>
<td>3</td>
<td>35.0</td>
<td>35.237</td>
</tr>
<tr>
<td>4</td>
<td>25.0</td>
<td>24.856</td>
</tr>
<tr>
<td>6</td>
<td>12.0</td>
<td>12.340</td>
</tr>
<tr>
<td>8</td>
<td>6.2</td>
<td>6.091</td>
</tr>
<tr>
<td>10</td>
<td>3.1</td>
<td>2.972</td>
</tr>
</tbody>
</table>

Table 3 Predicted Values from Estimation of Parameters

The default values are used in the following sample program to produce Figure 1.

```sas
GOPTIONS DEVICE=C MGOF97L GSFMODE=REPLACE GSNAME=GSASFILE RESET=GLOBAL;
SYMBOL1 C=RED V=NONE I=J;
SYMBOL2 V=DOT C=BLUE I=NONE;
SYMBOL3 C=GREEN V=NONE I=J;
```
TIT1: USING THICKER LINES

You can use the 'WIDTH' statement to control the thickness of the lines. 'WIDTH=n' statement specifies the width (in positive integer pixels) of the line. The WIDTH statements are added to the following sample code.

symbol1  c=red v=none i=j width=20;
symbol2 v=dot c=blue i=none width=16;

axis1 label = ( angle=90 c=blue 'Concentration(nmol/L/mg)') width=20;
axis2 length = 60 pct
  order = (-1 to 11 by 1)
  label = (c=blue 'Time(hrs)') width=20;
run;

TIP 2: USING BOLD FONTS

The fonts available with SAS/GRAPH software are listed in 'SAS/GRAPH Software Reference' document. These fonts are stored in the library SASHELP.FONTS.

goptions device=cgmof97l gsfmode=replace
gsfname=gsasfile reset=global ftext=zapfb

title c=red h=2.8 'Time vs. Plasma Concentration';

axis1 label = ( angle=90 c=blue h=2.3
  'Concentration(nmol/L/mg)') width=20
  value = ( c=black );

axis2 length = 60 pct
  order = (-1 to 11 by 1)
  label = (c=blue h=2.3
    'Time(hrs)') width=20
    value = ( c=black );
run;
TIP 3: COMBINING WITH OTHER APPLICATION SOFTWARE FOR ANNOTATION

The Annotate facility enables you to generate a data set of graphics command that can produce graphics output. With the Annotate facility, you create a data set that contains observations for every graphics element drawn. But creating an Annotate data set is a time consuming process.

The application software selected for this graphics output enhancement is Microsoft PowerPoint. You can import your graphics to PowerPoint and perform the following tasks:

- Use a full set of drawing tools for replacement of Annotate data set.
- Insert words and create multiple fonts and sizes for labels, titles, footnote, annotation, and legend.
- Combine words and images.
- Arrange and rearrange the graphics layout.

The following two figures, Figures 4 and 5, are graphics examples that combine PowerPoint for annotation.

TIP 4: USING COLORS FOR DATA GROUPING
A set of laboratory data collected from a clinical trial is used for the illustration purposes. The following program is used to produce Figure 6.

```plaintext
goptions device=cgmof97l gsfmode=replace
gsfname=gsasfile reset=global ftext=zapfb;
run;
symbol1 v=dot c=red i=none width=16;
symbol2 v=dot c=blue i=none width=16;
symbol3 c=black v=none i=j width=20;
symbol4 c=blue v=none i=j width=20;
symbol5 c=red v=none i=j width=20;
title c=red h=2.7 'FSH Comparison:
Baseline vs. Week 4';
axis1 label = ( angle=90 c=blue h=2.3
'Week 4') width=20
value  = ( c=black );
axis2  length = 60 pct
label  = (c=blue h=2.6 'Baseline')
width=20 value  = ( c=black );
proc gplot data = fad;
plot  week4a*basea=1 week4d*based=2
i*ii=3 aa*i=4 dd*i=5/ nolegend overlay
vaxis=axis1 haxis=axis2;
run;
```

Figure 6. Example of Using Colors for Data Grouping

TIP 5: ADDING RIGHT AXIS APPEARANCE

The AXIS statement allows you to control the appearance of the axes of your plots and charts. The following example uses both left and right axes appearance for easier references of data points. The following program uses multiple AXIS statements to control the right and left axes appearances.

```plaintext
goptions device=cgmof97l gsfmode=replace
gsfname=gsasfile reset=global ftext=zapfb;
run;
symbol1 v=dot c=red i=none width=16;
symbol2 v=dot c=blue i=none width=16;
symbol3 c=black v=none i=j width=20;
title c=red h=2.7 'FSH Comparison:
Baseline vs. Week 4';
axis1 label = ( angle=90 c=blue h=2.3
'Week 4') width=20
order = (0 to 140 by 10)
value  = ( c=black );
axis2 label = none
order = (0 to 140 by 10)
value  = ( c=black );
axis3 length = 60 pct
label  = (c=blue h=2.6 'Baseline')
order = (0 to 140 by 10)
value  = ( c=black );
proc gplot data = fad;
plot  week4a*basea=1 week4d*based=2
i*ii=3 aa*i=4 dd*i=5/ nolegend overlay
vaxis=axis1 haxis=axis2;
run;
```

Figure 7. Example of Using Colors for Data Grouping
TIP 6: OVERLAYING THE OUTPUT FROM DIFFERENT GRAPHICS PROCEDURES

In some cases, you have trend or distribution curves that were produced from PROC GPLOT and output from PROC GCHART or other procedures. It is a challenge task to overlay the graphics output from different procedures. You can use the GREPLAY procedure to display multiple pictures on one page or per screen. Calling other application software, such as Microsoft PowerPoint, is the alternative way. The following sample program produces Figure 8.

```sas
options device=cgmof97l gsfmode=replace gsfname=gsasfile reset=global ftext=zapfb;
pattern1 c=green ;
pattern2 c=red ;
pattern3 c=brown ;
pattern4 c=orange ;
pattern5 c=blue ;
pattern6 c=pink ;
pattern7 c=yellow ;
title1 c=red h=2.2 'Distribution of Male Testosterone';
axis1 label = ( angle=90 c=blue h=2.3 'Number of Sample') width=20
   value = ( c=black );
axis2 length = 60 pct
   label = (c=blue h=2.6 'Testosterone(NG/DL)') width=20
   value = ( c=black );
proc gchart data = p1;
   vbar cat  / discrete
      raxis = axis1
coutline=black
      maxis = axis2
sumvar = scount
   space=3
      patternid=midpoint
;run;
```

The sample SAS statements for producing graphic chart shown in Figure 9 are as follows:

```sas
symbol1 c=black i=splines width=55;
proc gplot data=p2;
   plot scount*cat=1/noaxis noframe;
run;
```

Figure 8. Output of GCHART

Figure 9 is a transparent graphics. You can overlay both Figures 8 and 9 together to produce Figure 10.
CONCLUSIONS

The graphics output is an important means to convey the information. The enhancements to produce a clear and elegant presentation can be achieved by modifying the SAS statements or mixing with other application software.

This paper provides the following tips to enhance your graphics output.

* Uses 'WIDTH' statement to control the thickness of the lines.
* Uses bold fonts for graphics text.
* Combines with other application software.
* Uses colors for data grouping.
* Adds right axis appearance for data points references.

REFERENCES


TRADEMARKS

SAS is a registered trademark of SAS Institute Inc., in the USA and other countries.

Microsoft and PowerPoint are registered trademarks of Microsoft Corporation, in the USA and other countries.

© indicates USA registration.

AUTHOR CONTACT INFORMATION

Shi-Tao Yeh, Ph. D.
(610)787-3856(W)
E-mail: shi-tao_yeh-1@gsk.com