SAS System

Computing for Research I

01/24/2012 N. Baker

Introduction to SAS/GRAPH

- SAS/GRAPH is the primary graphics component of SAS system.
- Includes charts, plots, and maps in both 2 and 3 dimensions.
- Procedures included GCHART, GPLOT, GMAP, GCONTOUR etc...
- We will focus on PROC GPLOT

Examples What Can Be Done using SAS GRAPH

Waiting longer to tie the knot

Americans are waiting longer to get married than they did decades ago, with larget percentages of both men and women in their 20s and 30s who have yet to say "I do".



Percentage of men and women never married, by age...

AGE	MEN	1970	WOMEN	MEN	2003	WOMEN
20-24	54.7		35.8	86.0		75.4%
25-29	19.1		10.5	54.6		40.3
30-34	9.4		6.2	33.1		22.7
35-44	6.7		5.2	19.5		13.2





SAS/GRAPH Mandelbrot Plot



Introduction Elements of SAS/GRAPH

Elements of SAS/GRAPH Overview



Elements of SAS/GRAPH PROC GPLOT: Specifying an input data set

Similar to all other SAS PROC's

– Proc gplot data=<libname>.<data set><options>;

*using default Work library;
proc gplot data=twovar;

*setting new library; libname indata "c:\sasdata\datafiles"; proc gplot data=indata.twovars;

```
*specifying data to use
libname indata "c:\sasdata\datafiles";
proc gplot data=indata.twovar(where=(x<1));</pre>
```

Options include setting annotate data sets, image mapping for drill-down plots in web applications, Creating Uniform axis across plots, and specifying SAS catalog for placement of output.

Elements of SAS/GRAPH PROC GPLOT: Specifying an input data set

Similar to all other SAS PROC's

– Proc gplot data=<libname>.<data set><options>;

*using default Work library;
proc gplot data=twovar;

*setting new library; libname indata "c:\sasdata\datafiles"; proc gplot data=indata.twovars;

*specifying data to use
libname indata "c:\sasdata\datafiles";
 proc gplot data=indata.twovar(where=(x<1));</pre>

Options include setting annotate data sets, image mapping for drill-down plots in web applications, Creating Uniform axis across plots, and specifying SAS catalog for placement of output.

Elements of SAS/GRAPH PROC GPLOT: *Plotting*

- You can use up to 2 plots statements at a time, however, at least one Plot statement is required.
- The plot statement is used to control the axis, plotting points, labels, tick marks, and the plot legend.
- The only required arguments are...
 Plot < Y Variable>*<X Variable> / <options>;

Elements of SAS/GRAPH PROC GPLOT: Plotting Options

- Options for plotting
 - Plot options
 - Legend= or nolegend: specifies figure legend options
 - Overlay: allows overlay of more than one Y variable
 - Skipmiss: breaks the plotting line where Y values are missing
 - Appearance option
 - Axis: Specifies axis label and value options
 - Symbol: Specified symbol options
 - href, vref: Draws vertical or horizontal reference lines on plot
 - frame/fr or noframe/nofr: specifies whether or not to frame the plot
 - caxis/ca, cframe/cfr, chref/ch, cvref/cv, ctext/c: specifies colors used for axis, frame, text or reference lines.

Introduction to SAS/GRAPH

• We will begin with rather simple code and let SAS decide how our graph will look.

 Then we will step through a few options that allow us to control and adjust the graphic output.

Examples 2 Variable Plotting / Scatter plots

- Suppose subjects are given a doses of experimental medication based on body weight over a 24 hour period (mg/24hrs). Variable X
- On the following day, each subject had their Vascular Cell Adhesion Molecule (µg/ml) levels measured. Variable Y1
- The investigators are interested in seeing a plot of the dose given vs. the plasma VCAM levels to see if there may be an effect of the drug dose.

Very basic plot, below we get all of the default options. Not very exciting. Definitely not publication quality.



Very basic plot, below we get all of the default options. Not very exciting. Definitely not publication quality.



Examples

2 Variables: AXIS Statements

- AXIS<1..99> <options>;
 - Label Option;
 - Angle/a=degrees (0-359)
 - Color/c=text color
 - Font/f=font
 - Height/h=text height (default=1)
 - Justify=(left/center/right)
 - Label="text string"
 - Options precede label
- axis1 label=(a=90 c=black f="arial" h=1.2 "time" a=90 c=black f="arial" h=1.0 "hours");

Examples 2 Variables: AXIS Statements

- AXIS<1..99> <options>;
 - Order Option
 - Order=(a to b by c): major tick marks will show up at intervals based on c.
 - Example order=(0 to 3 by 1);
 - Value Option
 - value=("" ""): applies text label to each major tick.
 - Example Value=("Start" "Middle" "End")

Examples 2 Variables: AXIS Statements

Resets previous _____ goptions reset=global ; options Horizontal axis ----- axis1 label=(f='arial/bo' h=1.9 "Dose" justify=c f='arial/bo' h=1.3 "mg/24 Hrs"); (X Variable) Vertical axis ____→ axis2 label=(a=90 f='arial/bo' h=1.9 "Plasma Level"); (Y Variable) proc gplot data=twovar; plot y1*x / haxis=axis1 vaxis=axis2; run; Call Axis statements

NOTE: you can also place the AXIS statements within the gplot proc

Examples 2 Variables: AXIS Statements



mg/24 Hrs

Examples 2 Variables: AXIS Statement

goptions reset=global ;



Examples 2 Variables: AXIS Statement



Examples 2 Variables: AXIS Statement

goptions reset=global;



```
plot y1*x / haxis=axis1 vaxis=axis2;
```

```
run;
```



Plasma Level

mg/24 Hrs

- Symbol<1...255> <options>;
 - Symbol options
 - Color= value color
 - Ci=line color
 - Height=symbol height
 - Line=line type
 - Value=symbol
 - Width=thickness factor
 - Interpol=point interpolations

- Symbol<1...255> <options>;
 - Symbol options
 - Interpolation options
 - Join, box, hilo interpolation, regression, spline, standard deviations.
 - value options
 - Dot, circle, star, square, plus, minus, "text value".
 - Color options
 - 256 colors available,

Symbol options

- Interpolation options
 - None
 - Join: points connected by straight line
 - Needle: vertical line from horizontal axis to point
 - Stepx: (L,R,C) step function, stepxJ will add a verticle line to each step plot
 - stdkxxx: (M,P,J,B,T) k=1,2,3 (standard deviations) or
 - » stdM=SEM, stdp=uses pooled sample variance, stdj=joins the errors, T will give tops and bottoms to error lines, where B will request error bars.
 - HILOxxx: (T,B,C,J)

Symbol options

- Interpolation options
 - R-series interpolation
 - Rxxxxxxx
 - » RL: linear regression
 - » RQ: Quadratic Regression
 - » RC: Cubic Regression
 - » CLM: CI for mean predicted values
 - » CLI: CI for Individual predicted values
 - » 90, 95, 99: confidence limits
 - » Example: RLCLM95 -> Gives a linear regression line with the 95% CL for mean predicted values

Examples 2 Variables: SYMBOL Statement

symbol1 value=dot color=black interpol=none;

```
proc gplot data=twovar;
    plot y1*x / haxis=axis1 vaxis=axis2;
```





mg/24 Hrs

Examples 2 Variables: Adding Regression Lines



Regression Equation: y1 = 0.481173 + 1.269433*x

Examples Grouping Variables

Many times we want to look at group differences.

Demographic groups, treatment groups, etc...

• Grouping variable must be in the data file.

Examples Grouping Variables

```
goptions reset=global ;
```

```
axis2 label=(a=90 f='arial/bo' h=1.9 "Plasma Level")
        order=(0 to 3 by 1)
        value=(a=90 f='arial' h=1.3 "0.0" "1.0" "2.0" "3.0");
```

You need to add a new SYMBOL statement for the each additional group.

Add the grouping variable to the PLOT statement

```
symbol1 value=dot color=black interpol=none;
symbol2 value=triangle color=black interpol=none;
```

Examples Grouping Variables



Examples Grouping Variables: Legend Statement

- Legend<1...99> <options>;
 - Legend options
 - Across=: number of columns
 - Down=: number of rows
 - Frame/noframe
 - Position=(bottom, middle, top) (left, center, right) (inside, outside)
 - Origin=(x,y)
 - Label=
 - Order=
 - Value=

These options are the same as within the axis statement discussed earlier

Examples Grouping Variables: Legend Statement

```
goptions reset=global ;
                 axis1 label=(f='arial/bo' h=1.9 "Dose" justify=c
                                     f='arial/bo' h=1.3 "mg/24 Hrs" )
                          order=(0 to 2 by 0.5)
                          value=(f='arial' h=1.3 "0.0" "0.5" "1.0" "1.5" "2.0");
                 axis2 label=(a=90 f='arial/bo' h=1.9 "Plasma Level")
                          order=(0 \text{ to } 3 \text{ by } 1)
                          value=(a=90 f='arial' h=1.3 "0.0" "1.0" "2.0" "3.0");
                 symbol1 value=dot color=black interpol=none(h=1.2
                 symbol2 value=triangle color=black interpol=none h=1.5;
                 legend1 across=1 down=2 noframe
Legend
                              position=(bottom right inside) mode=protect
Statement
                              label=(f='arial/bo' h=1.4 "Gender" )
                              value=(f='Arial/bo' h=1.4 "Female" "Male");
                 proc gplot data=twovar;
                       plot y1*x=gender / haxis=axis1 vaxis=axis2 legend=legend1;
                 run;
Call Legend
Statement
```

Examples Grouping Variables: Legend Statement



- Suppose that you have many observations on each subject taken at various time points.
- 40 subjects
- 2 treatments (Placebo and Active med)
- 5 time points (baseline plus 4 1-week intervals)
 - During the last week, both treatment groups receive Placebo
- Data should be in the Long format

At diagnosis, subjects are randomized to an experimental treatment or placebo. During the final week of treatment, all subjects will receive active medication.

Create appropriate axis and legend statements as before.

AXIS for X axis1 label=(f="arial/bo" h=1.5 "Time Since Diagnosis: Weeks") (time) variable order=(1 to 5 by 1) value=(f="arial" h=1.2 "Baseline" "1" "2" "3" "4") offset=(1,1);AXIS for Y axis2 label=(f="arial/bo" h=1.5 a=90 "Response") (Response) order=(0 to 100 by 10) value=(f="arial" h=1.2 "0" "10" "20" "30" "40" "50" variable "60" "70" "80" "90" "100") offset=(1,1); legend1 label=(f="arial" h=1.3 "Treatment Group") value=(f="arial" h=1.2 "Treatmant A" "Placebo") Added TITLE position=(top left inside) statement for mode=protect noframe; plot title "Individual Disease Progression":





Individual Disease Progression





Time Since Diagnosis: Weeks



Plot data by trt group and create a symbol statement for each group



Individual Disease Progression



Examples Using the Overlay statement to stack plots

Suppose that you are asked to graphically show progression of tumor growth for a group of subjects and overlay the progression of each treatment group.

50 subjects randomized to either low or high dose medication.

- Tumor size is measured at baseline as well as the following 9 weeks.
- The investigator would like an easy to present plot containing both pieces of information for a presentation to his peers.



Time Since Diagnosis: Weeks

```
axisl label=(f="arial/bo" h=1.5 "Time Since Randomization: Weeks")
order=(1 to 10 by 1)
value=(f="arial" h=1.2 "Baseline" "1" "2" "3" "4" "5" "6" "7" "8" "9")
offset=(1,1);
```

axis2 label=(f="arial/bo" h=1.5 a=90 "Tumor Growth")
 order=(0 to 80 by 10)
 value=(f="arial" h=1.2 "0" "10" "20" "30" "40" "50" "60" "70" "80")
 offset=(1,1);

```
axis3 label=(f="arial/bo" h=1.5 a=90 "")
order=(0 to 80 by 10)
value=(f="arial" h=1.2 "0" "10" "20" "30" "40" "50" "60" "70" "80")
offset=(1,1);
```

```
legend1 label=(f="arial" h=1.3 "Treatment Group")
    value=(f="arial" h=1.2 "Low Dose" "High Dose" )
    position=(top left inside)
    mode=protect noframe;
```

title "Individual Disease Progression";

```
proc gplot data=overlay;
        plot y*time=id / nolegend haxis=axis1 vaxis=axis2;
        plot2 y*time=trt / overlay legend=legend1 vaxis=axis3;
        symbol1 c=black i=join r=50 w=0.5;
        symbol2 c=blue i=stdmj l=1 w=4;
        symbol3 c=red i=stdmj l=1 w=4;
run;
```

Individual Disease Progression



Tumor Growth

Time Since Randomization: Weeks

Examples

Overlay multiple plots from different variables

```
proc logistic data=analysis desc;
    where nephropathy ne .;
    model nephropathy = log_oxldl_chol_base/clodds=wald;
    units log oxldl chol base= SD;
    output out=OXresults p=predict l=lower u=upper xbeta=logit / alpha=0.05;
run;
proc gplot data=oxresults;
    plot predict*log_oxldl_chol_base
        lower*log_oxldl_chol_base
        upper*log_oxldl_chol_base
        /overlay vaxis=axis1 haxis=axis2 nolegend;
run;
```

Use proc logistic to output the predicted probability of developing nephropathy given the baseline Oxidized LDL immune complex level as well as the 95% confidence limits.

Many PROCs can output predicted values, adjusted means, along with point wise confidence values.

Examples Overlay multiple plots from different variables



Examples Overlay multiple plots from different variables



Examples

Overlay multiple plots from different variables

```
symboll v=none i=box00f c=white bwidth=1;
symbol2 v=none i=box00f co=libr cv=libr bwidth=10 w=6;
symbol3 v=none i=box00f co=black cv=black bwidth=0.5;
symbol5 v=none i=box00f co=librar cv=librar bwidth=10 w=6;
symbol6 v=none i=box00f co=black cv=black bwidth=0.5 ;
symbol7 v=none i=join l=1 c=vigb w=6;
axisl label=(f="arial/bo" h=1.9 "Baseline OxLDL-IC Quartile")
      order=(-0.5 to 3.5 by 0.25)
      value=(f="arial/bo" h=1.5 "" "" "lst" "" "" "2nd" "" "" "3rd" "" ""
            "" "4th" "" "" )
      offset=(1,1);
axis2 label=(f="arial/bo" h=1.9 a=90 "Baseline LDL & HDL (mg/dl)") minor=none
      order=(0 to 250 by 25)
      offset=(5 pct)
      value=(f="arial/bo" h=1.3 "0" "" "50" "" "100" "" "150" "" "200" "" "250")
axis3 label=(f="arial/bo" h=1.9 a=270 "Baseline OxLDL-IC") minor=none
      order=(0 to 500 by 50)
      offset=(5 pct)
      value=(f="arial/bo" h=1.3 "0" "" "100" "" "200" "" "300" "" "400" "" "500"
legend label=(f="Arial/BO" h=1.5 'Baseline Characteristics')
      position=(top left inside) across=3 mode=share noframe
      Value=(f="arial" h=1.4 "OxLDL-IC" "LDL Cholesterol" "" "HDL Cholesterol"
            "");
proc gplot data=plots;
     plot median*rank=group/ noframe haxis=axisl vaxis=axis2 legend=legend;
      plot2 Oxmean*rank / overlay noframe haxis=axis1 vaxis=axis3 legend=legend
                        skipmiss;
run;
quit; run;
```

The Annotate Facility allows SAS users to customize graphical output. The customizations can be data driven or user specified. Text, shapes, lines and images can be added to output graphics

Step 1. Create an annotate data set

This data file will give commands to SAS/GRAPH

Specific variables must be in the annotate data set. Others are allowed but ignored.

What, how, and where are defined by these variables.

Table 1 list important variables.

TABLE 1. ANNOTATE DATA SET VARIABLES		
VARIABLE	DESCRIPTION	
FUNCTION	Specifies the Annotate drawing action. Table 2 below gives a list of important	
	functions.	
Х	The numeric horizontal coordinate.	
Y	The numeric vertical coordinate.	
Z	For three-dimensional graphs, specifies the coordinate for the 3 rd dimension.	
HSYS	The type of units for the size (height) variable.	
XSYS	The coordinate system for the X variable.	
YSYS	The coordinate system for the Y variable.	
ZSYS	The coordinate system for the Z variable (for three-dimensional graphs).	
ANGLE	Angle of text label or start angle for a pie slice.	
COLOR	Color of graphics item.	
LINE	Line type of graphics item.	
POSITION	Placement/alignment of text.	
ROTATE	Angle of individual characters in a text string or the sweep of a pie slice.	
SIZE	Size of the graphics item. Specific to the function. For example, size is the	
	height of the character for a label function.	
STYLE	Font/pattern of a graphics item.	
TEXT	Text to use in a label, symbol, or comment.	
WHEN	Determines if Annotate command is executed (B)efore or (A)fter the graph.	

The Annotate FUNCTION variable tells SAS what to do

TABLE 2. FUNCTIONS	
FUNCTION	DESCRIPTION
LABEL	Draws text.
MOVE	Moves to a specific point.
DRAW	Draws a line from the current position to a specified position.
COMMENT	As a documentation aid, allows you to insert a comment into the SAS
	Annotate file.
POLY	Specifies the starting point of a polygon.
POLYCONT	Continues drawing the polygon.
BAR	Draws a rectangle from the current position to a specified position
SYMBOL	Draws a symbol.
PIE	Draws a pie slice, circle or arc.

The annotate coordinate system allows for flexibility in placing objects within the output. There are 12 possible conditions.

Area	<u>Unit</u>	Coordinat	e System
Data	% Values	Absolute 1 2	Relative 7 8
Graphics Output Area	% Cells	Absolute 3 4	Relative 9 A
Procedure Output Area	% Cells	Absolute 5 6	Relative B C

Table 3. ANNOTATE MACROS

MACRO	DESCRIPTION
%DCLANNO	Declares the Annotate variables.
%LABEL(x, y, text-string, color, angle, rotate, size,	Places a label of text .
style, position)	
%MOVE(x, y)	Moves to a location.
%DRAW(x, y, color, line, size)	Draws a line from the current location to the
	specified location.
%COMMENT(text-string)	Allows an unexecuted comment to be inserted into
	the Annotate data set.
%POLY(x, y, color, style, line)	Begins drawing a polygon.
%POLYCONT(x, y, color)	Continues drawing a polygon.
%BAR(x1, y1, x2, y2, color, line, style)	Draws a bar.
%LINE(x1, y1, x2, y2, color, line, size)	Draws a line.
%PIEXY(angle, size)	Draws a pie slice.
%CIRCLE(x, y, size, color)	Draws a circle.

Proc GPLOT global options help make graphs more pleasing, however, there are cases where more work is needed to fully explain the data

Mean HbA1c % durring DCCT/EDIC study





```
%annomac;
data anno bar;
      %dclanno; length text $30;
      xsys='2'; ysys='2'; hsys='2';
      when='A';
      %bar(1.021, 6, 2.935, 10, white, 3, solid);
      %bar(11.055, 6, 11.940, 10, white, 3, solid);
      %bar(2.935, 6, 11.02,9.6, CX808080, 3, r5);
      %bar(11.95, 6, 20.07,9.6, CX808080, 3, r5);
      function='label'; color='black'; x=1.9; y=9.1; style='ARIAL/bo';
            text='Intensive';output;
      function='label'; color='black'; x=1.9; y=9.0; style='ARIAL/bo';
            text='Treatment';output;
      function='label'; color='black'; x=1.9; y=7.3; style='ARIAL/bo';
            text='Standard';output;
      function='label'; color='black'; x=1.9; y=7.2; style='ARIAL/bo';
            text='Treatment';output;
      function='label'; color='black'; x=7; y=9.8; size=0.19; style='ARIAL/bo';
            text='DCCT Trial';output;
      function='label'; color='black'; x=16.5; y=9.8; size=0.19; style='ARIAL/bo';
            text='EDIC Follow Up';output;
run;
```



Anno data set 1: Will place the image of the dollar over the plotting area.

Anno data set 2: Will create white Space above the Plotted line over time.

SET the anno data sets and call them in the GPLOT statement

```
data annodollar;
 length function $8;
 xsys='2'; ysys='1'; when='b';
 function='move'; x=&minyear; y=.1; output;
 function='image'; x=&maxyear; imgpath='C:\Documents and
       Settings\nab42\Desktop\dollar.jpg'; style='fit'; y=99.6; output;
run;
data annoblock; set data;
 xsys='2'; x=year; when='b';
 ysys='2'; y=dollarvalue; function='move'; output;
 ysys='1'; y=99.5; function='draw'; color="&backcolor";output;
run;
data myanno; set annodollar annoblock;
run:
proc qplot data=data anno=myanno;
       plot dollarvalue*year / vaxis=axis1 haxis=axis2;
run:
quit;
```



Awesome SAS v9.3 Upgrades

•You no longer need to turn on/off ODS Graphics for modeling outputs (proc phreg, proc reg, proc model ect...).

•Using png outputs, SAS Graph now uses anti-aliasing to create smoother more publication quality lines.

•Fill and text colors on graphs can now be specified to be semitransparent using Alpha Channel Color Transparency.

•Plot markers such as "square" can now be filled with the specified marker color using the command "squarefilled" or "triangefilled".

The End



"Our statistician will drop in and explain why you have nothing to worry about."