Distribution curves graphic, with patterned areas between minimum and maximum ranges using SAS®.

Author: Sylvain Cadieux, Senior SAS programmer, Algorithme Pharma Inc., Laval, Qc, Canada

ABSTRACT
Distribution plot graphic with patterned area helps to rapidly show differences between two compared items. They are under documented for proc gplot since they are mainly used in pie-charts. Nevertheless, we found an innovative way to do it. This paper explains how to plot a distribution graphic comparing values of 2 items in graphic using patterned areas of different colors for each item. The difficulties encountered with the SAS-Graph® language to produce this graphic are discussed.

KEYWORDS
SAS Graph, program, gplot procedure, graphic, shaded area, minimum maximum range, annotate, data.

INTRODUCTION
Distribution plot graphic with patterned area helps to rapidly show differences between two compared items. An example is presented in Figure A.

![Concentration range distribution](image)

Two things are important when planning a graph with SAS.

1. Use the appropriate SAS code
2. Pre-process the data for the graph.
In order to do a plot graphic with pattern area covering the minimum-maximum range values (like figure A), four variables of data are needed: one for minimum and maximum of two compared drugs or metabolites as shown in table 1.

**Table 1**

<table>
<thead>
<tr>
<th>time1</th>
<th>timeH</th>
<th>minA</th>
<th>minB</th>
<th>maxA</th>
<th>maxB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>8</td>
<td>9</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>1.5</td>
<td>10</td>
<td>20</td>
<td>46</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>11</td>
<td>16</td>
<td>45</td>
<td>54</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>14</td>
<td>12</td>
<td>47</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>15</td>
<td>9</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>16</td>
<td>8</td>
<td>42</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>17</td>
<td>7</td>
<td>33</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>20</td>
<td>12</td>
<td>32</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>17</td>
<td>17</td>
<td>30</td>
<td>45</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>16</td>
<td>18</td>
<td>29</td>
<td>45</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>28</td>
<td>49</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>11</td>
<td>8</td>
<td>27</td>
<td>43</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>4</td>
<td>1</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>17</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>18</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Processing **Table 1** data with the program **graph-figure1.sas** would result in **figure B**.

![Figure B](image-url)
SAS-GRAPH DIFFICULTIES TO SOLVE

Trying to add patterned lines without data transformation would give a graphic similar to figure C. Note that when SAS is told to add a pattern under the curves, it tries to close the curve figures by returning to the point of origin (0,0). In addition, the legend does not reflect the different patterns.

Figure C

EXPLANATION

SAS graph's patterned areas are mainly designed for graphics like pie-charts or bar-charts but not really for kinds of graphics like plot charts. To introduce patterned areas in plot charts, it is needed to draw a closed area in order to make SAS understands the limits of the area. By doing so, SAS-Graph would draw patterned lines between the curves (like figure A).

SOLUTION

Some modifications are required on data and program to obtain the desired graphic. Note that the legend in figure A has only two series but appears to draw four curves. On the other hand the legends in figure B and figure C show four series of data. This is because minimum and maximum series for the same drug have been merged together to produce figure A. It is needed to concatenate the minimum and maximum values of each drug in order to obtain a set of continuous values to feed the graphic. If the area is not fully closed, some surprises may occur when trying to add "patterned lines" between minimum and maximum curves as shown in figure C.

REORGANIZE DATA

It is important to close the area to be shaded to prevent SAS-Graph to produce patterns similar as figure C. Data must be reorganized by sorting the minimum and maximum by time order, and add a second time variable for both min and max values with this exception: one variable (max) has to be sorted in ascending sequence order while the other one (min) must be sorted in descending sequence as a mirror image. Then by appending one to the other to obtain a single series of min-max, a line will start from the origin of the plot going to the maximum time point and then go back to the origin and close the area of the figure. This little trick will help to provide the expected results in the graph. This means to concatenate min and max in a single variable. Also the time values must go back and forth in order to obtain a closed area line and then SAS is able to fill the area with the patterned lines.
Figure D illustrates how the data Table 1 is transformed to obtain Table 2 data. maxA and maxB are transferred in the first part of new variable minmaxA and minmaxB while the contents of the variables minA and minB are transferred in the bottom part.

Table 1

<table>
<thead>
<tr>
<th>time1</th>
<th>timeH</th>
<th>minA</th>
<th>minB</th>
<th>maxA</th>
<th>maxB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>6</td>
<td>9</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>8</td>
<td>14</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>1.5</td>
<td>10</td>
<td>20</td>
<td>46</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>11</td>
<td>16</td>
<td>45</td>
<td>54</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>14</td>
<td>12</td>
<td>47</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>15</td>
<td>9</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>18</td>
<td>8</td>
<td>42</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>17</td>
<td>7</td>
<td>33</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>20</td>
<td>12</td>
<td>32</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>17</td>
<td>17</td>
<td>30</td>
<td>45</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>16</td>
<td>18</td>
<td>29</td>
<td>45</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>28</td>
<td>49</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>11</td>
<td>8</td>
<td>27</td>
<td>43</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>4</td>
<td>4</td>
<td>21</td>
<td>31</td>
</tr>
<tr>
<td>17</td>
<td>20</td>
<td>1</td>
<td>0</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>18</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>time1</th>
<th>timeH</th>
<th>timerev</th>
<th>minmaxA</th>
<th>minmaxB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>3</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>1.5</td>
<td>5</td>
<td>46</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>6</td>
<td>45</td>
<td>54</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>7</td>
<td>47</td>
<td>45</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>8</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>9</td>
<td>42</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>10</td>
<td>33</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>11</td>
<td>32</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>12</td>
<td>30</td>
<td>45</td>
</tr>
<tr>
<td>13</td>
<td>11</td>
<td>13</td>
<td>29</td>
<td>45</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
<td>14</td>
<td>28</td>
<td>49</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>15</td>
<td>27</td>
<td>43</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>16</td>
<td>21</td>
<td>31</td>
</tr>
<tr>
<td>17</td>
<td>20</td>
<td>17</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>18</td>
<td>28</td>
<td>18</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>19</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Figure D

The first part of the data goes from the origin of the graphic to the right edge. This is the data step where max values are transferred into the new data table (Table 2). This code creates a new variable for min and max called minmax.
for each item. It populates minmax variables with max, then min will be used to build the second part of the minmax variable in the following data step. It will also create a timerev variable containing the same value as in time1 for the first part, except that the value order of time sequence will be reversed. The other values will be dropped since they are not required to feed the graphic.

First part of data is built here.

```plaintext
data library.data1;
  set library.oridata;
  timerev=time1;
  minmaxa=maxa;
  minmaxb=maxb;
  drop mina minb maxa maxb;
run;
```

For the second step, the data are rearranged to make the graphic going from the right edge of graphic back to the origin. Notice the code used to populate timerev variable. In fact all values are still linked to their respective times however timerev variable is created using the backward sequence. The mirror-reflect part of the time variable is created at this point. Ntime macro variable is used to determine the length of the data array on the time axis.

```plaintext
%let ntime=19;
data library.data2b;
  set library.oridata;
  timerev=(&ntime-time1)+1;
  keep time1 timeh timerev;
run;
```

The data is sorted in reverse order using the timerev variable, in order to be able to build a mirror-reflect of the data obtained in first part.

```plaintext
proc sort data=library.data2b out=library.data2br;
  by timerev;
run;
```

This step builds the time reverse variable timerev and keeps only the complement part of data to the first part.

```plaintext
data library.data2c;
  set library.oridata;
  timerev=(&ntime-time1)+1;
  minmaxa=mina;
  minmaxb=minb;
  keep timerev minmaxa minmaxb;
run;
```

Sorting data in reverse order

```plaintext
proc sort data=library.data2c out=library.data2cr;
  by timerev;
run;
```

This data step merges the 2nd part new files together

```plaintext
data library.data2new;
  merge library.data2br(in=A) library.data2cr(in=B);
  by timerev;
run;
```

Concatenate part 1 and part 2

```plaintext
data library.graphdata;
  set library.data1 library.data2new;
run;
```

Once the data has been transformed to obtain table 2 data, a big part of the work has been done. However there are a few things left to modify from the original program in order to obtain the final results as expected.
ORIGINAL PROGRAM (GRAPH-FIGURE1.SAS)

**------ Start of program ------**;
%let Vrange=(0 to 65 by 5);
%let Hrange=(0 to 35 by 5);
%let title=Concentration range distribution;
%let stitle1=(Study ABC);
%let graph1=figureA.PDF;
%let drug1=Drug-1;
%let drug2=Drug-2;
%let color1=red;
%let color2=blue;

**------ file + library definition ------**;
LibName  Library   '';

**------ graphical options ------**;
gooptions reset=all
  papersize=letter
  ftitle=swiss ftext=swiss htitle=2 htext=1.1
  rotate=landscape rotate;

**-----------------------------**;
axis1 order=&vrange offset=(0,0) minor=none label=(a=90 r=0 "Concentration (ng/mL)"

axis2 label=('"Time (hours)"') order=&Hrange;
axis3 label=none;

**------ Legend -------**;
legend1 label=none value=(tick=1 "&drug1 min" tick=2 "&drug2 min" tick=3 "&drug1 max"
  tick=4 "&drug2 max")
  shape=symbol(38pt,7pt) across=2 position=(bottom center outside);

**------ Titles ------**;
title1 f=swissb h=1.5 "&title";
title2 f=swiss h=1.3 "&drug1 vs &drug2 &stitle1";

**------ Lines to plot ------**;
%let min=MINA*timeh MINB*timeh;
%let max=MAXA*timeh MAXB*timeh;

**------ Graphic ------**;
ODS PDF file= "&graph1" BOOKMARKLIST=no;
ods graphics on;
proc gplot data=library.oridata;
  symbol1 interpol=join cv=&color1 co=&color1;
  symbol2 interpol=join cv=&color2 co=&color2;
  plot &min &max / overlay legend=legend1 vaxis=axis1 haxis=axis2;
run;
ods graphics off;
ODS PDF close;

**------ end of program ------**;

ADAPTATION OF THE PROGRAM

The legend need to be modified since the patterns do not appear (unlike pie-charts). However the proc annotate will be later used to fix this little problem. Prior to annotate, the legend will be modified to reflect the changes in data information (two series of data instead of four after the merge of minimum and maximum in a single variable).

Modification of the code in the legend section of the program:

legend1 label=none value=( tick=1 "&drug1 range(min-max)"
  tick=2 "&drug2 range(min-max)"
  shape=symbol(38pt,7pt) across=2 position=(bottom center outside);

Addition of extra code to the legend is needed, since SAS does not provide a box with pattern in the legend when performing a proc gplot. It is necessary to draw the boxes by annotating the graphic. A little trick here is to make believe SAS it is doing a gbar chart and have patterned boxes drawn over the area where a line appears.

Data library.anno1;
  length color $10 style $10;
  color="&COLOR1";
  line=0;
  size=2;
Modification to proc gplot, via annotate to reflect changes of data information in the program. Notice that there are now only two series of data.

CONCLUSION
With this little trick, and understanding SAS behavior, biostatisticians and SAS programmers can create graphic curves, with patterned areas between minimum and maximum ranges using SAS-graph and build even more sophisticated graphics adapted to their needs.

REFERENCES

ACKNOWLEDGEMENT
The author would like to thank Algorithm Pharma, for encouraging and supporting conference participation.

CONTACT INFORMATION
Your comments and questions are encouraged. Contact the author at:
Sylvain Cadieux,
Algorithm Pharma Inc.
Laval, Qc, Canada
Email: scadieux@algopharm.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.