Activity for Lab 8, ST512

- Recommendation: If students have time, they should type SAS code into the program editor, rather than copy and paste.

1. The data below are also available in the SAS program “germpct.sas” Write a datastep to read these data in as \( N = 12 \) observations at \( 3 \times 3 = 9 \) combinations, which are denoted with superscripts. Also, have SAS compute quadratic terms, \( moisture^2 \) and \( temp^2 \) and the product, \( moist \times temp \).

<table>
<thead>
<tr>
<th>Moisture</th>
<th>90</th>
<th>60</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp = 110</td>
<td>14(^1)</td>
<td>42(^2)</td>
<td>30(^3)</td>
</tr>
<tr>
<td>Temp = 90</td>
<td>42(^4)</td>
<td>85(^5), 92(^6), 84(^5), 78(^5)</td>
<td>51(^6)</td>
</tr>
<tr>
<td>Temp = 70</td>
<td>46(^7)</td>
<td>38(^8)</td>
<td>17(^9)</td>
</tr>
</tbody>
</table>

Your datastep might look like the following:

```sas
data germinate;
  input germpct moist temp trt @@;
  m2 = moist*moist; t2=temp*temp; mt = moist*temp;
cards;
  14  90  110 1  42  60  110 2  30  30  110 3
  42  90  90 4  85  60  90 5  51  30  90 6
    92  60  90 5
    84  60  90 5
    78  60  90 5
  46  90  70 7  38  60  70 8  17  30  70 9
;  
```

(Note the use of @@ to prevent SAS from moving to a new line to read the next record.)

2. Next, obtain a plot of germination percentage against moisture, with a different symbol for temperature (and vice-versa).

```sas
proc gplot data=germinate;
  plot germpct*moist=temp;
  plot germpct*temp=moist;
run;
```

Add a symbol of your liking:

```sas
symbol value=dot;
```

Better yet, have `proc gplot` fit a quadratic response, with different shapes for different levels of the factor other than the one on the horizontal axis:

```sas
symbol value=dot i=rq;  * i=rq means interpolate with Quadratic Regression;
```

3. Consider a 3d plot. To do this, create a grid of moisture and temperature, over which we'll obtain estimates of germination percentage by fitting the quadratic model below.

\[
\mu(x_M, x_T) = \beta_0 + \beta_{M1}x_M + \beta_{M2}x_M^2 + \beta_{T1}x_T + \beta_{T2}x_T^2 + \beta_{MT}x_Mx_T
\]

To construct a grid, which will be the floor of the plot, use nested do loops.
data grid;
  do moist = 30 to 90 by 2; do temp = 70 to 110 by 2;
    m2 = moist*moist; t2=temp*temp; mt = moist*temp;
    output; end; end;

Before merging the original germinate dataset with the grid dataset jitter the temperatures for the real data, so that they’ll be visible in the graph. Also, set a variable to enlarge the character size and color for real data (as opposed to that fitted over the grid):

data germinate;
  set germinate;
  *xtemp = temp+.0000*ranuni(1928766);* << no jittering;
  xtemp = temp+.0001*ranuni(1928766);* << imperceptible jittering;
  svar = 1.5 ; *<--- A size variable for graphing;
  cvar = 'green'; *<--- A color variable;
run;

data both;
  set grid germinate;
run;

4. Now fit the model again with proc reg, but this time a fitted values will produced for each of your grid values from the data in the grid dataset:

proc reg data=both;
  model germpct = moist temp t2 m2 mt;
  output out=out1 predicted = G_hat;
  Title "PROC REG to fit quadratic surface";
run;
data out1; set out1; if germpct = . then germpct = G_hat;
  if xtemp ne . then temp = xtemp; *<-- jittered version of temp;
  drop m2 t2 mt;
  ** Note germpct now has actual data followed by predictions over
    the moisture, temp grid we created;
proc print data=out1 (Obs = 30);run;

5. Plot the fitted values using the dataset you created to contain the fitted values:

proc g3d; scatter moist*temp=germpct / shape = "balloon" size = svar
  color = cvar noneedle;
  title "Observed data and fitted response surface";
run;

So far, so good?
6. Treating this as a one-factor experiment, with 9 levels of a single factor, obtain an ANOVA table. By comparing this ANOVA table to that from the quadratic fit above, obtain an F-ratio to test for lack-of-fit of the quadratic model, by hand.

   proc glm data=response; class trt; model germpct = trt;
   title "Getting pure error SS - can you obtain lack of fit F test?";
   title2 "Make sure the df make sense to you";

7. To get SAS to help, use proc rsreg.

   proc rsreg data=response;
   model germpct = moist temp;
   run;

   Now add the lackfit option after a slash in the model statement.

8. Discuss the output. At what combination of temperature and pressure is germination percentage maximized?

This activity has touched on several concepts:

- Factorial experiments
- Fancy 3-d plotting with proc g3d
- Response surface designs (and proc rsreg)
- Lack of fit