Luck and Foegeding data on caramel color

Jason A. Osborne, Dept. of Statistics, NCSU, Oct-Nov, 2002

Data to be analyzed were collected in two single-factor experiments. In one experiment (#1), three color responses (*lh ah bh*) were measured using four different corn syrups (26,42,55,62) and a fixed level of phospholipid (30). In the other (#2) experiment, the color measurements were taken at six different levels of a phospholipid (0,1,5,10,20,30) using corn syrup #42 Combining these two experiments leads to an incomplete two-factor design. The data were analyzed using a one-way model and questions about simple effects of corn syrup for fixed phospholipid and simple phospholipid effects for fixed corn syrup were addressed by calculating appropriate sums of squares among the nine treatment means. The one-way fixed effects model for any one of the three responses is then

$$Y_{ijk} = \mu + (\alpha\beta)_{ij} + E_{ijk}$$

where $j = 2$ for $i = 1, 2, \ldots, 5$ and $j = 1, 2, 3, 4$ for $i = 6$ for a total of nine treatment combinations. All nine treatment sample sizes are $n_{ij} \equiv 6$ with the exception of $n_{62} = 12$. since this treatment was observed in both experiments. In the model above, $k = 1, 2, \ldots, n_{ij}$ for each $ij$ combination. Additionally, a term can be added to the model to account for a possible experiment effect. This term would be estimated by considering the appropriate difference of sample means from the $i = 6, j = 2$ treatment combination, which is the only treatment combination with data from both experiments. This experiment effect is then nested in the treatment effect: let $z_{ij}$ be an indicator variable defined by

$$z_{ij} = \begin{cases} 0 & \text{expt \#1} \\ 1 & \text{expt \#2} \end{cases}$$

The model with an experiment effect is becomes

$$Y_{ijk} = \mu + (\alpha\beta)_{ij} + \beta z_{ij} + E_{ijk}.$$

To investigate whether there is evidence of an association between mean color and corn syrup, the null hypothesis

$$H_0 : (\alpha\beta)_{61} = (\alpha\beta)_{62} = (\alpha\beta)_{63} = (\alpha\beta)_{64} = 0$$

was formulated. The numerator sum of squares for an $F$-ratio to test this hypothesis is based on the first four treatment means:

$$\sum_{j=1}^{4} \sum_{k=1}^{n_i} (\bar{y}_{6j+} - \bar{y}_{6++})^2$$

obtained using the SLICE= option in an LSMEANS statement within PROC GLM in SAS. Similarly, to test for a simple phospholipid effect with the second corn syrup fixed, the following null hypothesis was formulated

$$H_0' : (\alpha\beta)_{12} = (\alpha\beta)_{22} = \cdots = (\alpha\beta)_{62} = 0.$$

The numerator sum of squares for the appropriate $F$-ratio is based on the following 6 treatment means:

$$\sum_{i=1}^{6}\sum_{k=1}^{n_i}(\bar{y}_{i2+} - \bar{y}_{+2+})^2.$$

To investigate a possible experiment effect, consider $H_0'' : \beta = 0$.

The table below indicates that for each response variable, the simple effect of corn syrup when phosholipid is 30 ($i = 6$) is highly significant, as is the simple effect of phospholipid when cornsyrup 42 ($j = 2$) is used.

| Response | $F$ and $p$ for $H_0$ | $df$ | $F$ and $p$ for $H_0'$ | $df$ | $F$ and $p$ for $H_0''$ |
|---|---|---|---|---|---|
| $lh$ | 73.24 ($p < .0001$) | 3 | 2736.97 ($p < .0001$) | 5 | 2.23 ($p = 0.1413$) |
| $ah$ | 102.86 ($p < .0001$) | 3 | 2652.79 ($p < .0001$) | 5 | 240.97 ($p < .0001$) |
| $bh$ | 18.76 ($p < .0001$) | 3 | 31.56 ($p < .0001$) | 5 | 3.90 ($p = 0.0538$) |

There was strong evidence of an experiment effect for the $ah$ response. To see this, note that the $bh$ sample means by experiment within the $i = 6, j = 2$ treatment combination of phospholipid = 30 and corn syrup = 42 are

$$\begin{aligned} \bar{y}^1_{62+} &= \tfrac{1}{6}\textstyle\sum_{k=1}^{6} y_{62k} &= 8.39 \quad \text{(expt \#1)} \\ \bar{y}^2_{62+} &= \tfrac{1}{6}\textstyle\sum_{k=7}^{12} y_{62k} &= 6.74 \quad \text{(expt \#2)} \end{aligned}$$

The difference is 1.65 with a standard error of 0.11, for a $t$ statistic of $t = 15.5$ which differs significantly from ($p < 0.0001$). (The $F$-ratio in the table above is just $241 = F = t^2 = 15.5^2$.)

The treatment means for the three responses are given below:

| LSMEAN NUMBER | Level of plipid | Level of cornsyrup | N | ----------lh Mean | -----------ah Mean | -----------bh Mean |
|---|---|---|---|---|---|---|
| 1 | 30 | 26 | 6 | 52.0866667 | 6.18333333 | 17.3833333 |
| 2 | 30 | 42 | 12 | 47.2150000 | 7.56666667 | 15.5425000 |
| 3 | 30 | 62 | 6 | 46.7466667 | 6.65833333 | 15.8066667 |
| 4 | 30 | 55 | 6 | 42.0516667 | 7.61666667 | 14.0250000 |
| 5 | 0 | 42 | 6 | 69.1883333 | -1.16000000 | 11.1800000 |
| 6 | 1 | 42 | 6 | 46.3800000 | 0.75333333 | 13.0533333 |
| 7 | 5 | 42 | 6 | 46.1750000 | 3.88166667 | 13.3616667 |
| 8 | 10 | 42 | 6 | 44.7616667 | 5.44500000 | 12.6500000 |
| 9 | 20 | 42 | 6 | 45.6466667 | 7.41000000 | 12.3166667 |

Subsequent multiple comparisons among the corn syrup means and among the phospholipid means for any of the three response variables can be made by inspecting the appropriate Bonferroni-adjusted pvalues in the matrices on the next pages. For example, to compare the mean $bh$ for corn syrups #26 and #62, which have LSMEAN numbers 1 and 3, we need to look at the pvalue in the $1^{st}$ row and $3^{rd}$ column of the matrix for $bh$. The difference 17.38-15.81=2.57

has a $p$ value, denoted w/ an asterisk, is 0.0350. Most relevant differences are statistically significant.

Lastly, normal plots and residual-by-predicted plots of residuals for the color measurements did not reveal any obvious asymmetry or inhomogeneity of variance.

```
            Least Squares Means for effect plipid*cornsyrup
                 Pr > |t| for H0: LSMean(i)=LSMean(j)
             Adjustment for Multiple Comparisons: Bonferroni

                       Dependent Variable: lh
```

| i/j | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | | <.0001 | <.0001 | <.0001 | <.0001 |
| 2 | <.0001 | | 1.0000 | <.0001 | <.0001 |
| 3 | <.0001 | 1.0000 | | <.0001 | <.0001 |
| 4 | <.0001 | <.0001 | <.0001 | | <.0001 |
| 5 | <.0001 | <.0001 | <.0001 | <.0001 | |
| 6 | <.0001 | 1.0000 | 1.0000 | <.0001 | <.0001 |
| 7 | <.0001 | 1.0000 | 1.0000 | <.0001 | <.0001 |
| 8 | <.0001 | 0.0042 | 0.1845 | 0.0076 | <.0001 |
| 9 | <.0001 | 0.3663 | 1.0000 | <.0001 | <.0001 |

| i/j | 6 | 7 | 8 | 9 |
|---|---|---|---|---|
| 1 | <.0001 | <.0001 | <.0001 | <.0001 |
| 2 | 1.0000 | 1.0000 | 0.0042 | 0.3663 |
| 3 | 1.0000 | 1.0000 | 0.1845 | 1.0000 |
| 4 | <.0001 | <.0001 | 0.0076 | <.0001 |
| 5 | <.0001 | <.0001 | <.0001 | <.0001 |
| 6 | | 1.0000 | 0.7491 | 1.0000 |
| 7 | 1.0000 | | 1.0000 | 1.0000 |
| 8 | 0.7491 | 1.0000 | | 1.0000 |
| 9 | 1.0000 | 1.0000 | 1.0000 | |

Dependent Variable: ah

| i/j | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | | <.0001 | 0.0016 | <.0001 | <.0001 |
| 2 | <.0001 | | <.0001 | 1.0000 | <.0001 |
| 3 | 0.0016 | <.0001 | | <.0001 | <.0001 |
| 4 | <.0001 | 1.0000 | <.0001 | | <.0001 |
| 5 | <.0001 | <.0001 | <.0001 | <.0001 | |
| 6 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |
| 7 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |
| 8 | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 |
| 9 | <.0001 | 1.0000 | <.0001 | 1.0000 | <.0001 |

| i/j | 6 | 7 | 8 | 9 |
|---|---|---|---|---|
| 1 | <.0001 | <.0001 | <.0001 | <.0001 |
| 2 | <.0001 | <.0001 | <.0001 | 1.0000 |
| 3 | <.0001 | <.0001 | <.0001 | <.0001 |
| 4 | <.0001 | <.0001 | <.0001 | 1.0000 |
| 5 | <.0001 | <.0001 | <.0001 | <.0001 |
| 6 | | <.0001 | <.0001 | <.0001 |
| 7 | <.0001 | | <.0001 | <.0001 |
| 8 | <.0001 | <.0001 | | <.0001 |
| 9 | <.0001 | <.0001 | <.0001 | |

Dependent Variable: bh

| i/j | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | | 0.0007 | 0.0350* | <.0001 | <.0001 |
| 2 | 0.0007 | | 1.0000 | 0.0105 | <.0001 |
| 3 | 0.0350 | 1.0000 | | 0.0085 | <.0001 |
| 4 | <.0001 | 0.0105 | 0.0085 | | <.0001 |
| 5 | <.0001 | <.0001 | <.0001 | <.0001 | |
| 6 | <.0001 | <.0001 | <.0001 | 1.0000 | 0.0044 |
| 7 | <.0001 | <.0001 | <.0001 | 1.0000 | 0.0004 |
| 8 | <.0001 | <.0001 | <.0001 | 0.1291 | 0.0706 |
| 9 | <.0001 | <.0001 | <.0001 | 0.0143 | 0.5296 |

| i/j | 6 | 7 | 8 | 9 |
|---|---|---|---|---|
| 1 | <.0001 | <.0001 | <.0001 | <.0001 |
| 2 | <.0001 | <.0001 | <.0001 | <.0001 |
| 3 | <.0001 | <.0001 | <.0001 | <.0001 |
| 4 | 1.0000 | 1.0000 | 0.1291 | 0.0143 |
| 5 | 0.0044 | 0.0004 | 0.0706 | 0.5296 |
| 6 | | 1.0000 | 1.0000 | 1.0000 |
| 7 | 1.0000 | | 1.0000 | 0.8738 |
| 8 | 1.0000 | 1.0000 | | 1.0000 |
| 9 | 1.0000 | 0.8738 | 1.0000 | |

```
*options ls=150;
options ls=75;

data one;
   input plipid  cornsyrup  rep  lh  ah  bh;
   *ah=sqrt(2+ah);
   *bh=exp(-0.5*log(bh));
   if _n_ <= 36 then expt=1;
   else expt=2;
   cards;
0  42  1  66.17  -1.11  12.18
0  42  2  70.96  -1.21  10.95
0  42  3  70.55  -1.19  10.56
0  42  4  66.17  -1.1  12.19
0  42  5  70.91  -1.2  11.03
0  42  6  70.37  -1.15  10.17
1  42  1  44.84  0.93  12.56
1  42  2  45.42  0.85  12.91
1  42  3  48.88  0.58  13.72
1  42  4  44.84  0.95  12.59
1  42  5  45.42  0.86  12.95
1  42  6  48.88  0.35  13.59
5  42  1  45.39  4.13  12.9
5  42  2  44.66  3.98  12.54
5  42  3  47.99  3.52  14.31
5  42  4  47.2  4.03  14.18
5  42  5  44.65  3.99  12.5
5  42  6  47.16  3.64  13.74
10  42  1  45.34  5.17  12.59
10  42  2  43.87  5.55  12.29
10  42  3  45.02  5.5  13.09
10  42  4  45.38  5.38  12.59
10  42  5  43.94  5.6  12.3
10  42  6  45.02  5.47  13.04
20  42  1  45.73  7.27  10
20  42  2  45.75  7.36  10.5
20  42  3  45.42  7.58  11
20  42  4  45.79  7.28  14.26
20  42  5  45.77  7.38  14.13
20  42  6  45.42  7.59  14.01
30  42  1  46.56  8.3  14.86
30  42  2  46.24  8.56  15.21
30  42  3  47.22  8.31  15.22
30  42  4  46.65  8.31  14.91
30  42  5  46.31  8.57  15.23
30  42  6  47.27  8.29  15.16
30  26  1  51.89  6.22  17.43
30  26  2  51.52  6.18  17.09
30  26  3  52.69  6.09  17.59
```

```
30  26  4  52.06  6.36  17.5
30  26  5  51.63  6.13  17.19
30  26  6  52.73  6.12  17.5
30  42  1  47.21  7.02  16
30  42  2  48.57  6.42  15.91
30  42  3  47.57  6.84  16.04
30  42  4  46.85  6.97  15.85
30  42  5  48.64  6.30  16.21
30  42  6  47.49  6.91  15.91
30  62  1  45.99  6.84  15.68
30  62  2  46.66  6.66  16.3
30  62  3  47.35  6.49  15.7
30  62  4  45.83  6.96  15.61
30  62  5  46.77  6.66  15.91
30  62  6  47.88  6.34  15.64
30  55  1  41.43  7.71  13.74
30  55  2  42.31  7.59  13.98
30  55  3  42.31  7.63  14.42
30  55  4  41.49  7.66  13.58
30  55  5  42.12  7.56  14.03
30  55  6  42.65  7.55  14.4
;
run;

proc sort;
   by descending expt;
run;
proc glm order=data;
   title "nested";
   class plipid cornsyrup expt;
   *model lh ah bh=plipid*cornsyrup expt/solution;
   model lh ah bh=plipid*cornsyrup expt(plipid*cornsyrup)/solution;
   lsmeans plipid*cornsyrup/slice=plipid pdiff adj=bon;
   lsmeans plipid*cornsyrup/slice=cornsyrup;
   lsmeans expt(plipid*cornsyrup);
   means expt(plipid*cornsyrup) plipid*cornsyrup;
   output out=two p=p1 p2 p3 r=r1 r2 r3;
run;

proc gplot;
   plot r1*p1=plipid;
   plot r2*p2=plipid;
   plot r3*p3=plipid;
   plot r1*p1=cornsyrup;
   plot r2*p2=cornsyrup;
   plot r3*p3=cornsyrup;
run;
proc univariate normal plot;
   var r1 r2 r3;
   histogram r1 r2 r3; run;
```