STAT502 Homework #6
due Monday, 8/7

1. (*adapted from Exercise 25.17*) The data in "pearls.txt" summarizes the market values for a sample of imitation pearls by number of coats of lacquer (Factor A) and randomly selected batch (Factor B). Twelve observations were recorded for each treatment combination. The first three columns correspond to market value, number of coats (6,8,10), and batch number, respectively. Use the following line to read the data into R:

```
data = read.table('pearls.txt', header=T)
```

   (a) Assuming the three coat levels are the only ones of interest, write the appropriate ANOVA model, including a term for AB interaction. Indicate which terms are random and which are fixed.

   (b) Construct a table summarizing the mean squares, degrees of freedom, and expected mean squares for this situation. Use the tables given in Lab 10 as a guide, but use numeric values where available from the data.

   (c) State the hypotheses for the test of interaction. Carry out this test with $\alpha = .05$.

   (d) Test for a main effect due to number of coats. State the hypotheses, and report the test statistic and conclusion with $\alpha = .05$.

   (e) Repeat part (b) for the main effect due to batches.

   (f) Estimate the difference in mean market values when comparing pearls with 6 coats to pearls with 8 coats (the first two Factor A groups). Also, estimate this difference when comparing groups 6 and 10 (the first and third Factor A groups). Use the Bonferroni adjustment so that both intervals can be interpreted simultaneously with 90% confidence.

2. (*adapted from R. Lymann Ott, 1993*) In a study of batch and site variability on blood pressure, a random sample of three batches was obtained from each of two randomly chosen blending sites. Five observations of content uniformity were then recorded for each batch.

   (a) Express the appropriate ANOVA model for Y_{ijk}, the kth uniformity observation from the jth bath in the ith site. Define any terms you use.

   (b) The following sum of squares were computed from this data:

$$\sum_i \sum_j \sum_k (\overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot\cdot\cdot})^2 = 0.01825$$

$$\sum_i \sum_j \sum_k (\overline{y}_{\cdot j\cdot} - \overline{y}_{\cdot\cdot\cdot})^2 = 0.01153$$

$$\sum_i \sum_j \sum_k (\overline{y}_{ij\cdot} - \overline{y}_{i\cdot\cdot} - \overline{y}_{\cdot j\cdot} + \overline{y}_{\cdot\cdot\cdot})^2 = 0.44249$$

$$\sum_i \sum_j \sum_k (y_{ijk} - \overline{y}_{ij\cdot})^2 = 0.29020$$

   Construct the ANOVA table for this situation with columns for source, SS, DF, MS, and expected MS. Note that the expected mean squares will be in terms of unknown parameters.

   (c) Using your table above, test for significant effects due to site and batch.

3. A study of plants in burned versus unburned areas involved repeated measurements of the same plots over a 15 month period. The floral count data is in the file 'floral.dat' and can be read into R with the commands

```
data = read.table('floral.dat',header=T,sep='\t')
trt = data$trt
plot = data$plot
time = data$time
y = data$resp
```

   (a) Provide an interaction plot of the data (with time on the horizontal axis), and comment on the effect of burning (versus not), the effect of time, and their interaction.

   (b) State the repeated measures ANOVA model, defining each term according to the description above. Also, indicate whether each term is fixed or random, and whether each term is nested or not.

   (c) What does this model assume about the covariance between observations from the same area and same plot? Does this seem appropriate here?

   (d) Fit the model above, and save the residuals. Does a histogram of the residuals support the assumption of normality of the error terms?

   (e) Carry out an appropriate test to see if the burned areas have significantly different floral counts when compared with the unburned areas, regardless of time.