Homework #6, ST518

1. A completely randomized experiment investigates the effects of increasing nitrogen (N) and copper (Cu) in the diet of chickens. Feed conversion ratio (FCR) is observed on $n = 4$ chickens for each of four treatment combinations(diets), with output below. Data are available online as "fcr.dat" so you can check your answers, but you should be able to complete these problems without software.

   (a) Write a factorial effects model for the 16 observed FCR measurements which assumes that, for a given diet, FCR is normally distributed, with variance σ^2 that is constant across diets.

   (b) Estimate the simple effect of increasing copper when $N = 25$.

   (c) Estimate the simple effect of increasing copper when $N = 45$

   (d) Estimate the difference in the simple effects of increasing copper across levels of Nitrogen.

   (e) Using significance level $\alpha = .05$, test the hypothesis that the simple effects of copper are constant across levels of nitrogen.

   (f) Report the smallest level of significance at which the difference between simple copper effects across levels of nitrogen may be declared significant.

   (g) Report a contrast sum of squares associated with the contrast tested in part (d).

   (h) Estimate the simple effect of increasing N when $Cu = 100$. Report a standard error and a 95% confidence interval for the effect. In light of this interval, can you declare the obserrved effect "significant" at level of significance $\alpha = .05$?

   (i) Estimate the main effect of increasing Cu. Give the F-ratio for a test of no effect, along with degrees of freedom.

   (j) Given the analysis you've done so far, is it appropriate to say the the effect of copper in this experiment is not significant (using level $\alpha = .05$)? Explain.

   (k) Report the contrast sums of squares for the main effect of copper and the main effect of nitrogen.

   (l) Obtain an ANOVA table which partitions the variability between the four treatments into meaningful components.

   (m) Briefly characterize the observed effects of copper and nitrogen on FCR, reporting appropriate p-values along the way.

   (n) It turns out that a control was also run (with $n = 4$), without any added Cu or N. The mean FCR was $\bar{y}_0 = 131$. The observed contrast of this mean with the average of the others is $\hat{\theta} = 131 - (1/4)(133 + 130 + 146 + 127) = -3$. Compute SS(diet), the diet sum of squares (on $df = 5 - 1 = 4$) from a one-way analysis of variance using all five diets.

```
The SAS System
The MEANS Procedure
                                    N
          Cu              N       Obs            Mean         Variance
-------------------------------------------------------------------
          10             25         4     133.0000000       36.0000000
                         45         4     130.0000000       16.0000000
         100             25         4     146.0000000       58.6666667
                         45         4     127.0000000       57.3333333
-------------------------------------------------------------------
```

2. Consider the barley sprouting data ("barley.dat") which come from a completely randomized, crossed design in which 30 lots of barley seeds are randomized to 10 combinations of amount of water and age of seeds prior to planting. The response is the number of seeds that sprout. (Each lot is 100 seeds.)

   (a) Posit a factorial effects model for these data. Why might the homogeneity of variance assumption might be questionable?

   (b) Report the p-value for a test of interaction between water and seed age.

   (c) What does it mean for the effects of water and seed age to be plausibly additive?

   (d) Obtain a 95% confidence interval for the effect of water. Does more water lead to more sprouting seeds?

   (e) Qualitatively assess the effect of seed age. Are older seeds more likely to sprout? (Alternatively, estimate the linear effect of age, $\widehat{\theta} = (-2, -1, 0, 1, 2)(\bar{y}_{1..}, \ldots, \bar{y}_{5..})'$ and test that $\theta = 0$.)

   (f) After averaging over to two levels of water, test for lack of fit of a model in which the effect of seed age on sprouting frequency is linear.

3. Consider the shrimp weight gain data ("shrimp.dat"). Let μ_{ijk} denote the mean weight gain at levels i, j and k of temperature, density and salinity, respectively. Consider the following contrasts:

$$\begin{aligned}
\theta_1 &= -\mu_{1\cdot 1} + 0\mu_{1\cdot 2} + \mu_{1\cdot 3} \\
\theta_2 &= -\mu_{2\cdot 1} + 0\mu_{2\cdot 2} + \mu_{2\cdot 3} \\
\theta_3 &= \mu_{1\cdot 1} - 2\mu_{1\cdot 2} + \mu_{1\cdot 3} \\
\theta_4 &= \mu_{2\cdot 1} - 2\mu_{2\cdot 2} + \mu_{2\cdot 3} \\
\theta_5 &= \theta_2 - \theta_1 \\
\theta_6 &= \theta_4 - \theta_3
\end{aligned}$$

   (a) Estimate each of these contrasts.

   (b) Identify which estimates differ significantly from 0. Use the results to characterize the interaction betweeen temperature and salinity.