Assignment 3 — due February 11, 2019

1. An experiment was conducted to assess how dietary supplements provided to female mice might affect the development of the nervous systems of their offspring. The experiment consisted of 6 different diets:

   1 standard mouse chow

   2 standard mouse chow plus Vitamin E

   3 standard mouse chow plus Vitamin C

   4 standard mouse chow plus Vitamin E and Vitamin C

   5 standard mouse chow plus a double dose of Vitamin C

   6 standard mouse chow plus Vitamin B6

   In total, there were 18 female mice. Diets were assigned at random, so that for each diet, there were 3 female mice. Each female was housed in a separate cage, and each female mouse was fed her assigned diet for the duration of her pregnancy. For the females in this study, the number of baby mice (sometimes called "pups") ranged from 6 to 10. For each female, 1 pup was sampled at random, sacrificed, and its optic nerves were examined under a microscope. A scoring system was used to designate the degree of development of the optic nerves. For each pup, this was a score between 0 and 100; the larger the number, the greater the degree of development. The data are in the file `mouse.txt`.

   (a) Construct side-by-side boxplots for the 6 groups and comment. You may show side-by-side dot plots instead, or overlay dot plots over your boxplots.

   (b) In general, is there evidence that dietary supplementation has an effect on the development of the optic nerve?

   (c) Does a residual plot suggest any problems?

   (d) Do this part "by hand." That is, for each method determine the relevant "yardstick" for comparing pairs of means, and make displays similar to that used in class.

   Perform a comparison of all possible pairs of means, separately, using Fisher's LSD method, Tukey's method, Hayter's method, and Scheffe's method and compare your results from these. Note: The `qtukey` function in R will be useful.

   (e) Is there evidence of a linear pattern in the response as a function of the amount of Vitamin C?

   (f) Here is a set of orthogonal contrasts that is of interest, where μ_i means the population mean of diet i:

   - $\mu_1 + \mu_2 - \mu_3 - \mu_4$
   - $\mu_1 - \mu_2 + \mu_3 - \mu_4$
   - $\mu_1 - \mu_2 - \mu_3 + \mu_4$

   In words, explain the meaning of each contrast, and test whether the contrast is equal to zero or not.

2. Consider the following summary of an unbalanced, completely randomized experiment:

Trt	1	2	3	4
$\bar{y}_{i\cdot}$	39.3	40.1	42.0	43.0
n_i	2	25	25	2

   For this experiment the F statistic is 4.90. Using $\alpha = 0.05$, perform Fisher's (protected) LSD to compare the means in this experiment. Note that for an unbalanced experiment, this basically amounts to doing individual T tests for the various pairs of means (after the initial F test).

   When you do so, an apparently illogical result occurs. What is it, and why does it occur?

3. (a) This is a continuation of problem 2 from assignment 2. Consider the design matrix $\mathbf{X}$ consisting of the columns [1 x1 x2 x3 x4]. In part (b), you essentially tried to fit the model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ with $\boldsymbol{\beta}' = (\mu, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$. Show in this case that $\mathbf{X}$ is of rank less than 5.

   (b) If $\mathbf{X}$ is not of full rank, then $\mathbf{X}'\mathbf{X}$ is singular. Some authors deal with a singular $\mathbf{X}'\mathbf{X}$ by defining a "generalized inverse." If $\mathbf{A}$ is a matrix (square or not) then $\mathbf{A}^-$ is said to be a generalized inverse of $\mathbf{A}$ if $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$. All matrices $\mathbf{A}$ have a generalized inverse whether they are singular or not.

   Find a generalized inverse of $\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$.

   (Hint: you can find one that is a 2×2 matrix with only 1 non-zero entry.)

   (c) If, in a regression, $\mathbf{X}'\mathbf{X}$ is singular, we can still define $\hat{\boldsymbol{\beta}}$ by using

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$$

   Consider the (simpler, but unbalanced) problem with $\mathbf{Y}' = (72, 36, 12, 48, 12, 36)$ and

$$\mathbf{X}\boldsymbol{\beta} = \begin{bmatrix} \mu + \alpha_1 & & \\ \mu + \alpha_1 & & \\ \mu + \alpha_1 & & \\ \mu & +\alpha_2 & \\ \mu & +\alpha_2 & \\ \mu & & +\alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

   Show that each of

$$G_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$G_2 = \begin{bmatrix} 1 & -1 & -1 & 0 \\ -1 & 4/3 & 1 & 0 \\ -1 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$G_3 = \frac{1}{18} \begin{bmatrix} 0 & 2 & 3 & 6 \\ 0 & 4 & -3 & -6 \\ 0 & -2 & 6 & -6 \\ 0 & -2 & -3 & 12 \end{bmatrix}$$

$$G_4 = \frac{1}{54} \begin{bmatrix} 17 & -11 & -8 & 1 \\ -11 & 23 & 2 & -7 \\ -8 & 2 & 26 & -10 \\ 1 & -7 & -10 & 35 \end{bmatrix}$$

   is a generalized inverse of $\mathbf{X}'\mathbf{X}$ and calculate $\hat{\boldsymbol{\beta}}$ for each G. Is $\hat{\boldsymbol{\beta}}$ the same for all four generalized inverses? Is the estimate of μ the same? What about $\mu + \alpha_i$ for each i? What about $\sum \alpha_i$?

   (d) Repeat the above with a different $\mathbf{Y}$ (you choose the value of $\mathbf{Y}$). What tentative conclusions can you reach about the characteristics of the $\hat{\boldsymbol{\beta}}$ vector that are invariant as you change $\mathbf{Y}$?

   (e) What's interesting about G_3 and G_4?

4. Consider the usual two-factor experiment randomized as a CRD, with factors A and B, and with model $Y_{ijl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijl}$ and with restrictions: $\sum_i \alpha_i = 0$, $\sum_j \beta_j = 0$, $\sum_i (\alpha\beta)_{ij} = 0$ for all j, and $\sum_j (\alpha\beta)_{ij} = 0$ for all i. Suppose each factor is at 2 levels, so $i = 1, 2$ and $j = 1, 2$ and suppose $l = 1, \ldots, n$. You know that you can express A_{main} and AB as contrasts in the group means: $\overline{ab}, \overline{a}, \overline{b}, \overline{1}$.

   What is the relationship between A_{main} and $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\widehat{(\alpha\beta)}_{ij}$?

   What is the relationship between AB and $\hat{\alpha}_i$, $\hat{\beta}_j$, and $\widehat{(\alpha\beta)}_{ij}$?