BIOSTATISTICS 667
Homework 2

1. (This problem looks long, but it'll probably take the least time to do). Consider the model: $Y_1, \cdots, Y_n$ are iid random variable distributed as normal(μ, σ^2). Derive the REML log-likelihood in this problem with two options for R; option 1: $R_i = Y_i - Y_1$; option 2: $R_i = Y_i - \bar{Y}$, where $\bar{Y} = \sum_{i=1}^{n} Y_i / n$. Do you obtain the same log-likelihood (additive constants do not matter).

   Hints and comments:
   In applying the REML formula below, work with all n observations and n residuals (do not drop one residual as was done in class for the "direct" computation). The $n \times n$ matrix Σ is simply $\sigma^2 I$, and I denotes the identity matrix. The $n \times 1$ matrix X is 1_n, a column of n ones. The vector R is $n \times 1$, *not* $(n-1) \times 1$.

   The REML log-likelihood is:

   $$l_{REML}(\theta; R) = -\frac{1}{2} \log |\Sigma| - \frac{1}{2} \log |X^\top \Sigma^{-1} X| - \frac{1}{2} Q(R, \Sigma, X),$$

   where

   $$Q(R, \Sigma, X) = R^\top \left\{ \Sigma^{-1} - \Sigma^{-1} X (X^\top \Sigma^{-1} X)^{-1} X^\top \Sigma^{-1} \right\} R.$$

   The following may or may not be needed. The matrix $J_{n \times n}$ of 1's is $1_n 1_n^\top$. Further, if a is an $n \times 1$ vector then $a^\top 1 = 1^\top a = \sum_{i=1}^{n} a_i$, $a^\top a = \sum_{i=1}^{n} a_i^2$ and $a^\top J a = a^\top 1 1^\top a = (a^\top 1)(1^\top a) = (\sum_{i=1}^{n} a_i)^2$. The determinant of a diagonal matrix is simply the product of the diagonal elements.

2. The textbook presents tests of linear, qudratic and cubic contrasts in Table 3.1, page 44. What does it mean to say that the linear contrast (the true population contrast, not the estimate) is zero? What does it mean to say that the qudratic contrast is zero? Provide a clear and explicit interpretation. Assume four equally-spaced observations (as in the textbook).

3. The program "tlc006.sas" fits a collection of models. You do not need to run the program in order to answer this question, but you may need to consult the SAS manual.

   (a) For each model give a short description of the mean structure and of the covariance structure. Give the description in English *without* using any mathematical symbols.

   (b) For model M_7 give a full description using mathematical notation (design matrix, parameters, $\cdots$).

   (c) For model M_7, interpret the coefficients of "t6" and "t6*active".

   (d) In model M_7, determine which parmeter(s) reflect a treatment effect, and explain why.