# Assignment # 4, Stat 8311, Fall 2006

## Reading

Chapter 4 of the notes

## Problems

The following problems are due on Monday, October 16, 2006, in class.

1. Consider the simple linear regression model:

$$\mathrm{E}(y_i) = \mu_i = \alpha + \beta x_i, i = 1, \ldots, n$$

with var$(y) = \sigma^2 I$, and not all the $x_i$ are equal.

   (a) Describe $\mathcal{E}$ and $\mathcal{E}^{\perp}$, where $\mathcal{E}$ = estimation space.

   (b) Find $P_{\mathcal{E}} Y$ and $Q_{\mathcal{E}} Y$, where $Y = (y_1, ...y_n)'$.

   (c) Give explicitly Var$(Y - \hat{\mu})$, the covariance matrix of the residuals.

   (d) How would the above structure change if the $x$s were all equal? In particular, characterize the estimable linear combinations of $\alpha$ and $\beta$.

2. Suppose $y \in \Re^2$, Var$(y) = \sigma^2 I$ and E$(y_i) = \beta_1 - \beta_2, i = 1, 2$.

   (a) Which of the following are estimable: $\beta_1,\ \beta_2,\ \beta_1 - \beta_2,\ \beta_1 + \beta_2$?

   (b) For the estimable functions, find the best linear unbiased estimators and their variances.

   (c) Now consider the following mean structure:

$$E(y_1) = (1 - \eta)\beta_1 - (1 + \eta)\beta_2$$

$$E(y_2) = (1 + \eta)\beta_1 - (1 - \eta)\beta_2$$

   where $|\eta|$ is a small number. Repeat parts (a) and (b). What happens to the variances as $\eta \to 0$?

3. Consider the one-way anova model, with the over-parameterized mean function

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

for $i = 1, 2, 3,\ j = 1, 2, 3,$ E$(\varepsilon_{ij}) = 0$; Var$(\varepsilon) = \sigma^2 I$. Here are some data:

| I | II | III |
|---|---|---|
| 6 | 8 | 8 |
| 2 | 5 | 4 |
| 4 | 8 | 6 |

   (a) If no restrictions are put of $\mu$ and the $\alpha_i$, give all least squares estimates of the parameter vector, and a description of all possible estimable functions (give the description both numerically and in general).

   (b) If $\alpha_3$ is dropped from the model, give the least squares estimate of the parameter vector, describe in words the meaning of each of the parameters, and give a description of all possible estimable functions.

   (c) If $\alpha_3$ is returned to the model, and the restriction $\sum \alpha_i = 0$ is added to the model, give the least squares estimate of the parameter vector, describe in words the meaning of each of the parameters, and give a description of all possible estimable functions.

4. A *two-pan balance* is a weighing device in which an object of unknown weight is put in one pan, and objects of known weight are put into the other pan until the pans balance. Four objects with unknown weights $A, B, C$ and $D$ are to be weighed on a two pan balance. In design D1, we make four measurements $y_1, \ldots, y_4$. $y_1$ is the amount on the right-hand pan which together with $D$ will balance $A, B$, and $C$ collectively on the left hand pan, so

$$y_1 = A + B + C - D + \varepsilon_1$$

where $\varepsilon_i$ is random measurement error. Similarly, $y_2, y_3$ and $y_4$ are obtained by putting $C, B$, and $A$, respectively on the right-hand pan:

$$\begin{aligned} y_2 &= A + B - C + D + \varepsilon_2 \\ y_3 &= A - B + C + D + \varepsilon_3 \\ y_4 &= -A + B + C + D + \varepsilon_4 \end{aligned}$$

Assume that the measurement error has zero mean, is uncorrelated from measurement to measurement and has variance $\sigma^2$.

(a) Find $\mathcal{E}$ and its dimension.

(b) Find the BLUE of $A$ and its variance, or show $A$ to be nonestimable.

(c) Find the BLUE of the total weight $W = A+B+C+D$ and its variance, or show it to be nonestimable.

(d) In design D2, we use one trial to weigh all the objects; the second trial to compare objects A and B to C and D; the third trial compares A to B and the last C to D. The design is as follows:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} +1 & +1 & +1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & 0 & 0 \\ 0 & 0 & -1 & +1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}$$

For what linear combinations of A, B, C, and D does design D2 yield a BLUE with smaller variance that that for design D1?

5. Let $y$ be an $n$-vector of random variables such that $\text{Var}(y) = \sigma^2 I$ and $\text{E}(Y) = \mu \in \mathcal{E} = \text{R}(J_n)$. Thus, $\mu = \beta J_n$ for some scalar $\beta$. Consider the estimator $\tilde{\beta} = \sum w_i y_i / \sum w_i$ for $\sum w_i \neq 0$, and $\tilde{\mu} = \tilde{\beta} J_n$. Is $\tilde{\mu}$ a projection of $y$? An orthogonal projection? Can any projection of $y$ onto $\text{R}(J_n)$ be written in the above form? Be sure to provide sufficient support for your response.

6. Artificial heart valves are made of a ceramic material. To test the durability of the ceramic, an accelerated life test is performed on the valve in which the valve is cycled rapidly in a flowing solution that simulates the human blood system. The valve is removed from the system periodically, and the maximum wear is measured. A reasonable model for this experiment is

$$y_i = \beta x_i + (\varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_i), i = 1, 2, ..., n$$

where $y_i$ is the wear at the $i$-th observation, $\beta$ is an unknown constant corresponding the wear rate, and $x_i$ is the known number of cycles (or the equivalent number of years of wear) up to the $i$-th observation. The $\varepsilon_i$ are assumed independent random variables with mean 0 and variance $\text{var}(\varepsilon_i) = (x_i - x_{i-1})\sigma^2$ (set $x_0 = 0$).

(a) Compute the BLUE $\hat{\beta}$ of $\beta$, and find its variance, and also find the estimate of $\sigma^2$. To complete this problem, you will probably need the following matrix result:

$$\begin{pmatrix} a_1 & a_1 & a_1 & \cdots & a_1 \\ a_1 & a_2 & a_2 & \cdots & a_2 \\ a_1 & a_2 & a_3 & \cdots & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_p \end{pmatrix}^{-1} =$$

$$\begin{pmatrix} \frac{1}{a_1}+\frac{1}{a_2-a_1} & -\frac{1}{a_2-a_1} & 0 & 0 & \cdots & 0 \\ -\frac{1}{a_2-a_1} & \frac{1}{a_3-a_2}+\frac{1}{a_2-a_1} & -\frac{1}{a_3-a_2} & 0 & \cdots & 0 \\ 0 & -\frac{1}{a_3-a_2} & \frac{1}{a_4-a_3}+\frac{1}{a_3-a_2} & -\frac{1}{a_4-a_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{a_p-a_{p-1}} \end{pmatrix}$$

This last matrix is called an *symmetric band matrix* as only the diagonal and first off-diagonals are nonzero, and all other elements are zero.

(b) Derive a test of the hypothesis:

NH: $\beta = \beta_0$
AH: $\beta \neq \beta_0$

where $\beta_0$ is a fixed, predetermined number. Give the test statistic, and its distribution under the assumption of normality, for both the null and alternative hypotheses.