# LONDON TAUGHT COURSE CENTRE

## LTCC Basic Course — Statistical Modelling and Estimation

**Exercise Sheet 5: Solutions** — **February/March 2012**

---

1. Under the assumptions of and using the notation in Theorem 3.6 show that the test statistic $F$ in the theorem can be written as

$$F = \frac{(\mathbf{Y} - \mathbf{X}_0\hat{\gamma})'(\mathbf{Y} - \mathbf{X}_0\hat{\gamma}) - (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{\text{MSE}\,\text{rank}(\mathbf{M} - \mathbf{M}_0)},$$

where $\hat{\boldsymbol{\beta}}$ and $\hat{\gamma}$ are least squares estimates of the corresponding parameter vectors.

The numerator of the test statistic $F$ in Theorem 3.6 or (3.8) is equal to $\mathbf{Y}'(\mathbf{M}-\mathbf{M}_0)\mathbf{Y}$. Since $\mathbf{M} - \mathbf{M}_0 = (\mathbf{I} - \mathbf{M}_0) - (\mathbf{I} - \mathbf{M})$ this can be written as

$$\mathbf{Y}'((\mathbf{I} - \mathbf{M}_0) - (\mathbf{I} - \mathbf{M}))\mathbf{Y} = \mathbf{Y}'(\mathbf{I} - \mathbf{M}_0)\mathbf{Y} - \mathbf{Y}'(\mathbf{I} - \mathbf{M})\mathbf{Y}$$

The matrices $\mathbf{I} - \mathbf{M}_0$ and $\mathbf{I} - \mathbf{M}$ are idempotent and symmetric, which implies that $\mathbf{Y}'(\mathbf{I} - \mathbf{M}_0)\mathbf{Y} = ((\mathbf{I} - \mathbf{M}_0)\mathbf{Y})'(\mathbf{I} - \mathbf{M}_0)\mathbf{Y}$ and $\mathbf{Y}'(\mathbf{I} - \mathbf{M})\mathbf{Y} = ((\mathbf{I} - \mathbf{M})\mathbf{Y})'(\mathbf{I} - \mathbf{M})\mathbf{Y}$. Further by Theorem 2.2 we have that $\mathbf{M}_0\mathbf{Y} = \mathbf{X}_0\hat{\gamma}$ and $\mathbf{M}\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}}$, where $\hat{\gamma}$ and $\hat{\boldsymbol{\beta}}$ are least squares estimates of the parameter vectors $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$. Thus

$$\mathbf{Y}'(\mathbf{M} - \mathbf{M}_0)\mathbf{Y} = (\mathbf{Y} - \mathbf{X}_0\hat{\gamma})'(\mathbf{Y} - \mathbf{X}_0\hat{\gamma}) - (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}),$$

and so $F$ in Theorem 3.6 can be represented as claimed.

2. In the linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $V(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$ suppose that the vector $\boldsymbol{\Lambda}'\boldsymbol{\beta}$ is estimable so that $\boldsymbol{\Lambda}' = \mathbf{A}'\mathbf{X}$ for some matrix $\mathbf{A}$. Let $\mathbf{M}$ be the orthogonal projection matrix onto $C(\mathbf{X})$ and $\tilde{\mathbf{M}}$ be the orthogonal projection matrix onto $C(\mathbf{MA})$.

   (a) Show that for every vector $\mathbf{b}$ of appropriate dimension $\mathbf{b}'\mathbf{A}'\mathbf{X} = \mathbf{0}$ if and only if $\mathbf{b}'\mathbf{A}'\mathbf{M} = \mathbf{0}$.

   Suppose that $\mathbf{b}'\mathbf{A}'\mathbf{X} = \mathbf{0}$. Since $\mathbf{M} = \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ it then follows that $\mathbf{b}'\mathbf{A}'\mathbf{M} = \mathbf{b}'\mathbf{A}'\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}' = \mathbf{0}$. For the converse suppose that $\mathbf{b}'\mathbf{A}'\mathbf{M} = \mathbf{0}$ and note that $\mathbf{X} = \mathbf{MX}$. It then follows that $\mathbf{b}'\mathbf{A}'\mathbf{X} = \mathbf{b}'\mathbf{A}'\mathbf{MX} = \mathbf{0}$.

   (b) Hence prove that $\text{rank}(\boldsymbol{\Lambda}) = \text{rank}(\tilde{\mathbf{M}})$.

   The column spaces of the matrices $\tilde{\mathbf{M}}$ and $\mathbf{MA}$ are equal, that is $C(\tilde{\mathbf{M}}) = C(\mathbf{MA})$. It then follows from part (a) that

$$C(\boldsymbol{\Lambda}')^\perp = C(\mathbf{A}'\mathbf{X})^\perp = C(\mathbf{A}'\mathbf{M})^\perp = C(\mathbf{A}'\mathbf{M}')^\perp = C((\mathbf{MA})')^\perp$$

   and so $C(\boldsymbol{\Lambda}')^\perp$ and $C((\mathbf{MA})')^\perp$ have the same dimension. Let $m$ be the number of row of $\boldsymbol{\Lambda}'$. It then follows that $\dim(C(\boldsymbol{\Lambda}')^\perp) = m - \dim(C(\boldsymbol{\Lambda}')) = m - \text{rank}(\boldsymbol{\Lambda}')$ and $\dim(C((\mathbf{MA})')^\perp) = m - \dim(C((\mathbf{MA})')) = m - \text{rank}((\mathbf{MA})')$ and so $\text{rank}(\boldsymbol{\Lambda}') = \text{rank}((\mathbf{MA})')$. Because a matrix and its transpose have the same rank and since $C(\mathbf{MA}) = C(\tilde{\mathbf{M}})$ it finally follows that $\text{rank}(\boldsymbol{\Lambda}) = \text{rank}(\tilde{\mathbf{M}})$.

3. Consider the multiple regression model

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \epsilon_i$$

for three continuous explanatory variables $X_1, X_2, X_3$, where it is assumed that $\epsilon_i \sim N(0, \sigma^2)$ and that the random errors $\epsilon_i$ and $\epsilon_j$ are independent for $i \neq j$.

(a) For the data in the following table

| $Y$ | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|
| 10 | −1 | −1 | −1.0001 |
| 12 | −1 | 1 | −0.9999 |
| 20 | 1 | −1 | 1.0000 |
| 22 | 1 | 1 | 1.0000 |
| 16 | 0 | 0 | 0.0000 |
| 18 | 0 | 0 | 0.0000 |

test the hypothesis $\beta_1 = \beta_2 = 0$ against a two-sided alternative at the 5% level of significance.

This is equivalent to testing if the model can be reduced to the model

$$Y_i = \beta_0 + \beta_3 x_{i3} + \epsilon_i.$$

For this test we can use the $F$ statistic in Theorem 3.6. In the lectures it was noted that the form (3.10) of $F$ is frequently more convenient for hand calculations and hence we use this here, that is we calculate

$$F = \frac{(\mathbf{Y} - \mathbf{X}_0\hat{\boldsymbol{\gamma}})'(\mathbf{Y} - \mathbf{X}_0\hat{\boldsymbol{\gamma}}) - (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{\text{MSE}\,\text{rank}(\mathbf{M} - \mathbf{M}_0)},$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & -1 & -1.0001 \\ 1 & -1 & 1 & -0.9999 \\ 1 & 1 & -1 & 1.0000 \\ 1 & 1 & 1 & 1.0000 \\ 1 & 0 & 0 & 0.0000 \\ 1 & 0 & 0 & 0.0000 \end{bmatrix} \quad \text{and} \quad \mathbf{X}_0 = \begin{bmatrix} 1 & -1.0001 \\ 1 & -0.9999 \\ 1 & 1.0000 \\ 1 & 1.0000 \\ 1 & 0.0000 \\ 1 & 0.0000 \end{bmatrix}$$

and $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ are respectively the least squares estimates in the full and the reduced model. Note that $\text{rank}(\mathbf{M} - \mathbf{M}_0)$ in the denominator is equal $\text{rank}(\mathbf{X}) - \text{rank}(\mathbf{X}_0)$, as was explained in Section 3.3 of the lecture notes.

In the full model we have $\hat{\boldsymbol{\beta}} = [16.33\ 5\ 1\ 0]'$ and so the vector $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ of fitted values and the residual vector $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$ are equal to

$$\hat{\mathbf{y}} = \begin{bmatrix} 10.33 \\ 12.33 \\ 20.33 \\ 22.33 \\ 16.33 \\ 16.33 \end{bmatrix} \quad \text{and} \quad \mathbf{e} = \begin{bmatrix} -0.33 \\ -0.33 \\ -0.33 \\ -0.33 \\ -0.33 \\ 1.67 \end{bmatrix}$$

giving rise to a residual sum of squares equal to

$$\mathbf{e}'\mathbf{e} = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = 3.33.$$

Furthermore the mean square error in the full model is

$$\text{MSE} = \frac{(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{n - \text{rank}(\mathbf{X})} = \frac{3.33}{2} = 1.67.$$

In a similar way the least squares estimate of $\hat{\boldsymbol{\gamma}} = [\beta_0\ \beta_3]'$ in the reduced model is $\hat{\gamma} = [16.33\ 5]'$ and the vectors of fitted values $\hat{\mathbf{y}}_0 = \mathbf{X}_0\hat{\gamma}$ and residuals $\mathbf{e}_0 = \mathbf{y} - \hat{\mathbf{y}}_0$ are

$$\hat{\mathbf{y}}_0 = \begin{bmatrix} 11.33 \\ 11.33 \\ 21.33 \\ 21.33 \\ 16.33 \\ 16.33 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_0 = \begin{bmatrix} -1.33 \\ 0.67 \\ -1.33 \\ 0.67 \\ -0.33 \\ 1.67 \end{bmatrix}$$

with a corresponding residual sum of squares of

$$\mathbf{e}_0'\mathbf{e}_0 = (\mathbf{Y} - \mathbf{X}_0\hat{\gamma})'(\mathbf{Y} - \mathbf{X}_0\hat{\gamma}) = 7.33.$$

From these the observed value of the test statistic is

$$F = \frac{7.33 - 3.33}{1.67 \times 2} = 1.20.$$

Under $H_0$ the test statistic $F$ has an $F_{2,2}$ distribution. For the test at the 5% level we need the upper 5% point of this distribution, which can be found, for example, in the New Cambridge Statistical Tables to be equal to 19.00. Thus there is no evidence to reject $H_0$ at the 5% level of significance. You may also check that the p-value of the test is equal to 0.455.

Note that equivalently we could have used the test statistic in Section 3.4 of the lecture notes with some appropriate matrix $\boldsymbol{\Lambda}'$.

(b) For the data in (a) also test the hypothesis $\beta_1 = \beta_3 = 0$ against a two-sided alternative, again using a 5% significance level.

In the same way as in (a) we find that

$$F = \frac{103.33 - 3.33}{1.67 \times 2} = 29.94.$$

The critical value of the test is the same as in (a) and since $F > 19.00$ we reject $H_0$ at the 5% level of significance. The p-value can be found to be equal to 0.03.

4. The two-way analysis of variance model without interaction for two factors A and B with respectively $a$ and $b$ levels where each combination of the levels of A and B is replicated $r$ times is given by

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk},$$

where $i = 1, \ldots, a$, $j = 1, \ldots, b$ and $k = 1, \ldots, r$. Further it is assumed that the random errors $\epsilon_{ijk}$ are independent and identically distributed as $N(0, \sigma^2)$.

In what follows suppose that $a = b = 2$ and also that $r = 2$.

(a) Write the model in matrix notation with parameter vector $\boldsymbol{\beta} = [\mu\ \alpha_1\ \alpha_2\ \beta_1\ \beta_2]'$.
For $a = b = r = 2$ the model can be written in matrix notation as $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where

$$\mathbf{Y} = \begin{bmatrix} Y_{111} \\ Y_{112} \\ Y_{121} \\ Y_{122} \\ Y_{211} \\ Y_{212} \\ Y_{221} \\ Y_{222} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \end{bmatrix}.$$

It is further assumed that $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and $V(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$, which is equivalent to assuming that $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ and $V(\mathbf{Y}) = \sigma^2\mathbf{I}$.

(b) Find the test statistic $F$ for testing $H_0 : \boldsymbol{\Lambda}'\boldsymbol{\beta} = \mathbf{0}$, where

$$\boldsymbol{\Lambda}' = \begin{bmatrix} 1 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

and $\boldsymbol{\beta} = [\mu\ \alpha_1\ \alpha_2\ \beta_1\ \beta_2]'$. Also state the distribution of $F$ under $H_0$.
The vector $\boldsymbol{\Lambda}'\boldsymbol{\beta}$ is estimable since $\boldsymbol{\Lambda}' = \mathbf{A}'\mathbf{X}$ where

$$\mathbf{A}' = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

The $F$ statistic (3.12) for testing $H_0 : \boldsymbol{\Lambda}'\boldsymbol{\beta} = \mathbf{0}$ from Section 3.4 is

$$F = \frac{(\boldsymbol{\Lambda}'\hat{\boldsymbol{\beta}})'(\boldsymbol{\Lambda}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{\Lambda})^-\boldsymbol{\Lambda}'\hat{\boldsymbol{\beta}}}{\text{MSE}\,\text{rank}(\boldsymbol{\Lambda})},$$

where MSE is the mean square error from Section 2.2 of the lecture notes.
The rank of $\mathbf{X}$ is equal to 3. By using the first method in Section 1.3 with

$$\mathbf{X}^* = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

a g-inverse of $\mathbf{X}'\mathbf{X}$ can be found as

$$(\mathbf{X}'\mathbf{X})^- = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 \\ -1 & 2 & 2 & -1 & -1 \\ -1 & 2 & \frac{5}{2} & -\frac{5}{4} & -\frac{5}{4} \\ 0 & -1 & -\frac{5}{4} & \frac{11}{8} & \frac{9}{8} \\ 0 & -1 & -\frac{5}{4} & \frac{9}{8} & \frac{11}{8} \end{bmatrix}.$$

It follows that

$$\boldsymbol{\Lambda}'(\mathbf{X}'\mathbf{X})^-\boldsymbol{\Lambda} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}.$$

From Section 2.2 of the lecture notes the unique least squares estimate of $\mathbf{\Lambda}'\boldsymbol{\beta}$ is

$$\mathbf{\Lambda}'\hat{\boldsymbol{\beta}} = \mathbf{\Lambda}'(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y} = \frac{1}{4}\left[\begin{array}{c} Y_{111}+Y_{112}+Y_{121}+Y_{122} \\ Y_{211}+Y_{212}+Y_{221}+Y_{222} \end{array}\right] = \left[\begin{array}{c} \bar{Y}_{1\cdot\cdot} \\ \bar{Y}_{2\cdot\cdot} \end{array}\right]$$

where $\bar{Y}_{1\cdot\cdot}$ is the mean of the responses at level 1 of the first factor and $\bar{Y}_{2\cdot\cdot}$ is the mean of the responses at level 2 of the first factor.
The test statistic for testing $H_0 : \mathbf{\Lambda}'\boldsymbol{\beta} = \mathbf{0}$ is thus

$$F = \frac{2(\bar{Y}_{1\cdot\cdot}^2 + \bar{Y}_{2\cdot\cdot}^2)}{\text{MSE}},$$

where we have used that $\text{rank}(\mathbf{\Lambda}) = 2$. Under $H_0 : \mathbf{\Lambda}'\boldsymbol{\beta} = \mathbf{0}$ this has an $F_{2,5}$ distribution.