# 1 Exercises 4a p.103

3. **Done** If $\hat{\lambda}_H$ is the least squares estimate of the Lagrange multiplier associated with the constraints $A\beta = c$ show that

$$RSS_H - RSS = \sigma^2 \hat{\lambda}_H^T var(\hat{\lambda}_H)^{-1} \hat{\lambda}_H$$

*Proof*
From section 3.8.1 we know that

$$\hat{\lambda}_H = 2\left[A(X'X)^{-1}A'\right]^{-1}(A\hat{\beta} - c)$$

Then we can calculate $var(\hat{\lambda}_H)^{-1}$.

$$\begin{aligned}
var(\hat{\lambda}_H)^{-1} &= var(2\left[A(X'X)^{-1}A'\right]^{-1}(A\hat{\beta} - c))^{-1} && (1)\\
&= var(2\left[A(X'X)^{-1}A'\right]^{-1}A\hat{\beta})^{-1} && (2)\\
&= \frac{1}{4}\left[A(X'X)^{-1}A'\right]A^{-1}var(\hat{\beta})^{-1}\left(\left[A(X'X)^{-1}A'\right]A^{-1}\right)' && (3)\\
\varphi &= \frac{1}{4\sigma^2}\left[A(X'X)^{-1}A'\right]\left(A(X'X)^{-1}A'\right)^{-1}\left[A(X'X)^{-1}A'\right] && (4)
\end{aligned}$$

Then our new expression is

$$\begin{aligned}
RSS_H - RSS &= 4\sigma^2(A\hat{\beta} - c)\left[A(X'X)^{-1}A'\right]^{-1}\varphi\left[A(X'X)^{-1}A'\right]^{-1}(A\hat{\beta} - c) && (5)\\
&= (A\hat{\beta} - c)\left(A(X'X)^{-1}\right)^{-1}(A\hat{\beta} - c)' && (6)\\
& && (7)
\end{aligned}$$

So that we have

$$RSS_H - RSS = \sigma^2 \hat{\lambda}_H^T var(\hat{\lambda}_H)^{-1} \hat{\lambda}_H \qquad (8)$$

5. **Wrong** Consider the full rank model $X\beta = (X_1, X_2)(\beta_1', \beta_2')'$ where $X \sim n \times q$.

   (a) Obtain a test statistic for $\beta_2 = 0$.
   Let

   $$RSS_H - RSS = (A\hat{\beta} - c)\left(A(X'X)^{-1}A'\right)^{-1}(A\hat{\beta} - c)' \qquad (9)$$

   Then we need to find the matrix $A$ for the hypothesis $H_0 : \beta_2 = 0$. Then we want to let $c = 0$ from (9) and the for the matrix $A$ we have

   $$\begin{aligned}
   A\beta &= A_1\beta_1 + A_2\beta_2 && (10)\\
   &= A_2\beta_2 = 0 && (11)
   \end{aligned}$$

   Since $\beta_2 \sim q \times 1$, and for our particular $H_0$, we have $A_1 = 0$ and $A_2 \sim (q-1) \times q = I_{(q-1)\times q}$. Then we have

   (b) Find $E[RSS_H - RSS]$.

## 2 Exercises 4b p.109

1. **Done** Let $Y_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} + \varepsilon_i$, $i = 1, 2, \ldots, n$, where the $\varepsilon_i$ are independent $N(0, \sigma^2)$. Prove that the $F$-statistic for testing the hypothesis $H : \beta_r = \beta_{r+1} = \cdots = \beta_{p-1} = 0$ $(0 < r \le p-1)$ is unchanged if a constant, $c$, is subtracted from each $Y_i$.

   *Proof*
   Since $r > 0$, we could rewrite the model

   $$Y_i - c = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} \tag{12}$$

   as

   $$\begin{aligned} Y_i &= (\beta_0 + c) + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} & (13) \\ Y_i &= \beta_0^* + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} & (14) \end{aligned}$$

   and we are done.

2. **Done** Let $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $(i = 0, 1, \cdots, n)$. where $\varepsilon_i \sim N(0, \sigma^2)$.

   (a) Show that the correlation coefficient of $\hat{\beta}_0$ and $\hat{\beta}_1$ is $-n\bar{x}/(\sqrt{n \sum x_i^2})$.

   *Proof*
   We know that $var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$. Where

   $$\begin{aligned} (X'X)^{-1} &= \frac{1}{\sum (x_i - \bar{x})^2} \begin{pmatrix} \frac{1}{n} \sum x_i^2 & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} & (15) \\ \rho_{\hat{\beta}_0, \hat{\beta}_1} &= \frac{cov(\hat{\beta}_0, \hat{\beta}_1)}{\sqrt{var(\hat{\beta}_0) var(\hat{\beta}_1)}} & (16) \end{aligned}$$

   Then if

   $$\begin{aligned} \gamma &= \sum_{i=0}^{n} (x_i - \bar{x})^2 & (17) \\ & & (18) \end{aligned}$$

   We have

   $$\begin{aligned} \rho &= \frac{\sigma^2 (-\bar{x})/\gamma}{\sqrt{\sigma^4 (\frac{1}{n} \sum x_i^2)/\gamma)(1/\gamma)}} & (19) \\ &= \frac{-\bar{x}}{\sqrt{(\frac{1}{n} \sum x_i^2)}} & (20) \\ &= \frac{-n\bar{x}}{\sqrt{n \sum x_i^2)}} & (21) \end{aligned}$$

   and we are done.

(b) Derive an $F$-statistic for testing $H : \beta_0 = 0$.

*Proof*
Here we will use equation (9) with the following definitions:

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 \end{pmatrix} && (22) \\ c &= 0 && (23) \end{aligned}$$

Then the $F$-statistic is defined as

$$F = \frac{\hat{\beta}_0 \sum x_i^2}{S^2 \sum (x_i - \bar{x})^2} \tag{24}$$

3. **Partial** Given $\bar{x} = 0$, derive an $F$-statistic for $H : \beta_0 = \beta_1$. And show it is equivalent to a certain $t$-test.

Let $X \sim n \times 2$, $A \sim 1 \times 2$ where $A = \begin{pmatrix} 1 & -1 \end{pmatrix}, c = 0$. The $F$-statistic will be defined as

$$F = \frac{RSS_H - RSS/q}{RSS/(n-p)} \tag{25}$$

Where $RSS_H - RSS$ is defined as in (9). And we have

$$\begin{aligned} (X'X)^{-1} &= \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum x_i^2} \end{pmatrix} && (26) \\ A(X'X)^{-1}A' &= \frac{1}{n} + \frac{1}{\sum x_i^2} = \frac{\sum x_i^2 + n}{n \sum x_i^2} && (27) \\ A\hat{\beta} - c &= \hat{\beta}_0 - \hat{\beta}_1 && (28) \\ \frac{RSS}{n-p} &= S^2 && (29) \end{aligned}$$

Then the $F$-statistic is

$$F = \frac{(\hat{\beta}_0 - \hat{\beta}_1)^2 n \sum x_i^2}{S^2 (\sum x_i^2 + n)} \tag{30}$$

4. **Done** Let

$$Y = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ -1 & 1 \end{pmatrix} \theta + \varepsilon \tag{31}$$

And find an $F$-statistic for $H : \theta_1 = 2\theta_2$

Then we take the $F$-statistic as defined in (25) with $A = \begin{pmatrix} 1 & -2 \end{pmatrix}$ and $c = 0$. Where

$$\begin{aligned} (A\hat{\theta} - c) &= \theta_1 - 2\theta_2 && (32) \\ (X'X)^{-1} &= \begin{pmatrix} 1/2 & 0 \\ 0 & 1/6 \end{pmatrix} && (33) \\ A(X'X)^{-1}A &= 7/6 && (34) \\ F &= \frac{(\hat{\theta}_1 - 2\hat{\theta}_2)^2}{(7/6)S^2} && (35) \end{aligned}$$

5. **Done (Ryan)** Given $Y = \theta + \varepsilon$, where $\varepsilon \sim N(0, \sigma^2 I_4)$ and $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 0$, show that the $F$-statistic for testing $H : \theta_1 = \theta_3$ is

$$\frac{2(Y_1 - Y_3)^2}{(Y_1 + Y_2 + Y_3 + Y_4)^2}$$

According to Dr. Horvath, what they mean by $\mathbf{Y} = \boldsymbol{\theta} + \boldsymbol{\varepsilon}$ is the following:

$$\begin{aligned} Y_1 &= \theta_1 + \varepsilon_1 \\ Y_2 &= \theta_2 + \varepsilon_2 \\ Y_3 &= \theta_3 + \varepsilon_3 \\ Y_4 &= \theta_4 + \varepsilon_4 \end{aligned}$$

But with the restriction $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 0$ we can write it as:

$$\begin{aligned} Y_1 &= \theta_1 + \varepsilon_1 \\ Y_2 &= \theta_2 + \varepsilon_2 \\ Y_3 &= \theta_3 + \varepsilon_3 \\ Y_4 &= -(\theta_1 + \theta_2 + \theta_3) + \varepsilon_4 \end{aligned}$$

So we can write the model as $\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$, where:

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

Now with $\mathbf{A} = (1, 0, -1)$ you can find the $F$ statistic for the test and you'll get the answer in the book. I didn't use the Lagrange multiplier hint (or at least I didn't realize I was using it if I did), but doing it this way requires inverting a 3x3 matrix and I don't know if he'll ask one that requires that.

## 3 Exercises 4c p.113

1. **Done** Suppose that $\beta_1 = \cdots = \beta_{p-1} = 0$. Find the distribution of $R^2$ and hence prove that $E[R^2] = (p-1)/(n-1)$.

From example 4.8 in the book, we know that under $H : \beta_1 = \cdots = \beta_{p-1} = 0$, we have the $F$-statistic

$$F = \frac{R^2}{1-R^2}\frac{n-p}{p-1}$$

Which we rewrite

$$R^2 = \frac{\frac{p-1}{n-p}F}{1+\frac{p-1}{n-p}F} \tag{36}$$

Which is distributed like $Beta((p-1)/2, (n-p)/2))$, and has mean $(p-1)/(n-1)$.

2. **Partial,Most** For the general linear full-rank regression model, prove that $R^2$ and the $F$-statistic for testing $H : \beta_j = 0$ $(j \neq 0)$ are independent of the units in which the $Y_i$ and the $x_{ij}$ are measured.

   This means that we have the new model

$$Y/c = XK\beta + \varepsilon \tag{37}$$

   Where the matrix $K = diag(1, k_1, \ldots, k_{p-1})$ and we write

$$Y^* = X^*\beta^* + \varepsilon^* \tag{38}$$

   Then, to calculate the $F$-statistic we use (25) with the matrix $A \sim 1 \times p$ where the $j^{th}$ element is 1 and the rest are 0 and $c = 0$. We also note that the matrix $K$ is diagonal and, therefore, symmetric. Then we use the following definitions to calculate (25)

$$z_j = [(X'X)^{-1}]_{jj} \tag{39}$$

$$\begin{aligned}(A(X'^*X^*)^{-1}A')^{-1} &= (A(KX'XK)^{-1}A')^{-1} && (40)\\ &= (AK^{-1}(X'X)^{-1}K^{-1}A')^{-1} && (41)\\ &= \frac{k_j^2}{z_j} && (42)\end{aligned}$$

$$(A\hat{\beta}^* - c) = \frac{\hat{\beta}_j}{ck_j} \tag{43}$$

$$RSS_H - RSS/q = \frac{\hat{\beta}_j}{c^2 z_j} \tag{44}$$

$$\tag{45}$$

$$\begin{aligned}\hat{Y}^* &= XK(KX'XK)^{-1}KX'Y/c && (46)\\ &= XKK^{-1}(X'X)^{-1}K^{-1}KX'Y/c && (47)\\ &= X(X'X)^{-1}X'Y/c && (48)\\ &= \hat{Y}/c && (49)\end{aligned}$$

$$
\begin{aligned}
RSS &= (\hat{Y}^* - Y^*)'(\hat{Y}^* - Y^*) && (50)\\
&= \frac{RSS}{c^2} && (51)
\end{aligned}
$$

$$
\begin{aligned}
\frac{RSS_H - RSS/q}{RSS/(n-p)} &= \frac{\frac{\hat{\beta}_j}{c^2 z_j}}{RSS/(n-p)c^2} && (52)\\
&= \frac{\hat{\beta}_j}{z_j S^2} && (53)
\end{aligned}
$$

Which proves that the $F$-statistic is not affected by how the design points are measured.

# 4 Miscellaneous exercises 4 p.117

2. **Done** Given the two regression lines

$$Y_{ki} = \beta_k x_i + \varepsilon_{ki} \qquad (k = 1, 2; i = 1, 2, \ldots, n)$$

show that the $F$-statistic for testing $H : \beta_1 = \beta_2$ can be put in the form

$$F = \frac{(\hat{\beta}_1 - \hat{\beta}_2)^2}{2S^2(\sum x_i^2)^{-1}}$$

Obtain $RSS$ and $RSS_H$ and verify that

$$RSS_H - RSS = \frac{\sum x_i^2(\hat{\beta}_1 - \hat{\beta}_2)^2}{2}$$

First we define a combined model

$$Y = X\beta + \varepsilon$$

Where we define

$$Y = \begin{pmatrix} Y_{1,1} \\ Y_{1,2} \\ \vdots \\ Y_{1,n} \\ Y_{2,1} \\ Y_{2,2} \\ \vdots \\ Y_{2,n} \end{pmatrix} \tag{54}$$

$$X = \begin{pmatrix} x_1 & 0 \\ x_2 & 0 \\ \vdots & \vdots \\ x_n & 0 \\ 0 & x_1 \\ 0 & x_2 \\ \vdots & \vdots \\ 0 & x_n \end{pmatrix} \tag{55}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \tag{56}$$

Then we follow the regular procedure of using (25) to define the $F$-statistic. In this case, use $A = \begin{pmatrix} 1 & -1 \end{pmatrix}$ and $c = 0$. Then we have

$$(A(X'X)^{-1}A')^{-1} = \frac{\sum x_i}{2} \tag{57}$$

$$(A\hat{\beta} - c) = (\hat{\beta}_1 - \hat{\beta}_2) \tag{58}$$

Which gives us

$$F = \frac{\sum x_i(\hat{\beta}_1 - \hat{\beta}_2)}{2S^2} \tag{59}$$

And we are done since the second part of this question follows from the first.

4. **Done, Ryan** A series of $n+1$ observations $Y_i$ $(i = 1, 2, \ldots, n+1)$ are taken from a normal distribution with unknown variance $\sigma^2$. After the first $n$ observations it is suspected that there is a sudden change in the mean of the distribution. Derive a test statistic for testing the hypothesis that the $(n+1)^{st}$ observation has the same population mean as the previous observations.

   For this one, to set it up like a linear model, you can look at the model:

$$\begin{aligned} Y_1 &= \mu + \varepsilon_1 \\ Y_2 &= \mu + \varepsilon_2 \\ &\vdots \\ Y_n &= \mu + \varepsilon_3 \\ Y_{n+1} &= \lambda + \varepsilon_{n+1} \end{aligned}$$

Which we can write as $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where:

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\beta} = \begin{pmatrix} \mu \\ \lambda \end{pmatrix}$$

So we can test $H : \mu = \lambda$ using $\mathbf{A} = (1, -1)$. We get:

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} 1/n & 0 \\ 0 & 1 \end{pmatrix} \quad \text{so} \quad \hat{\beta} = \begin{pmatrix} \bar{Y}_n \\ Y_{n+1} \end{pmatrix} \quad \text{and} \quad A(\mathbf{X}^T\mathbf{X})^{-1}A^T = \frac{1}{n} + 1$$

And

$$RSS/(n-1) = \frac{1}{n-1}\left[\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \\ Y_{n+1} \end{pmatrix} - \begin{pmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_n \\ Y_{n+1} \end{pmatrix}\right]^T \left[\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \\ Y_{n+1} \end{pmatrix} - \begin{pmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_n \\ Y_{n+1} \end{pmatrix}\right] = S_n^2$$

So I get:

$$F = \frac{(\bar{Y}_n - Y_{n+1})^2}{S_n^2\left(1 + \frac{1}{n}\right)}$$

Now if you take the square root, it kind of looks like the answer in the back of the book. If you can get it into a t-distribution, let me know. It seems like it wouldn't be so bad, but I'm a little burned out.

## 5 Miscellaneous exercises 5 p.136

2. **Not Done** Prove that $(1 - \alpha/k)^k > 1 - \alpha$ $(k > 1)$.

6. **Done** Let $Y_I = \beta_0 + \beta_1 x_i + \varepsilon_i (i = 1, 2, \ldots, n)$, where the $\varepsilon_i$ are independently distributed as $N(0, \sigma^2)$. Obtain a set of multiple confidence intervals for all linear combinations $a_0\beta_0 + a_1\beta_1$ ($a_0, a_1$ not both zero) such that the overall confidence for the set is $100(1-\alpha)\%$.

   Use equation (5.22) from the book which gives a confidence interval as

   $$x'\hat{\beta} \pm (pF^{\alpha}_{p,n-p})^{1/2} S(x'(X'X)^{-1}x)^{1/2} \tag{60}$$

   with $x = (a_0, a_1)$ and recongizing that $(X'X)^{-1}$ can be defined as in (15). Then we get

   $$\begin{aligned} x'(X'X)^{-1}x &= \frac{1}{\varphi}(a_0^2 \sum(x_i^2/n) - 2a_0a_1\bar{x} + a_1^2 & (61) \\ \varphi &= \sum(x_i - \bar{x})^2 & (62) \end{aligned}$$

   And we are done recognizing that $x'\beta = a_0\beta_1 + a_1\beta_2$.

7. **Not Done** In constructing simultaneous confidence intervals for all $x'\beta$, explain why setting $x_0 \equiv 1$ does not affect the theory. What modifications to the theory are needed if $\beta_0 = 0$?