1. Assume that the following data come from the linear model:

\[ y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \]

\[ \varepsilon_i \sim N(0, \sigma^2) \quad i = 1, 2, \ldots, n \]

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>-6.1</td>
<td>-0.5</td>
<td>7.2</td>
<td>6.9</td>
<td>-0.2</td>
<td>-2.1</td>
</tr>
<tr>
<td>x</td>
<td>-2.0</td>
<td>0.6</td>
<td>1.4</td>
<td>1.3</td>
<td>0.0</td>
<td>-1.6</td>
</tr>
<tr>
<td></td>
<td>-1.7</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Find the maximum likelihood estimates of \( \beta_0 \), \( \beta_1 \), and \( \sigma^2 \)

\[
\hat{\beta} = (x^T x)^{-1} x^T y = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1.1630 \\ 3.2338 \end{bmatrix}
\]

\[
\hat{\sigma}^2 = \frac{1}{n} \hat{\varepsilon}^T \hat{\varepsilon} = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2 = \frac{24.0838}{8} = 3.01
\]

\[
s^2 = \frac{1}{n-k} \varepsilon^T \varepsilon = \frac{1}{n-k} \sum_{i=1}^{n} \varepsilon_i^2 = \frac{24.0838}{6} = 4.01
\]

2. The model:

\[ Y = \alpha_1 + \alpha_2 E_2 + \alpha_3 E_3 + u \]

is estimated by OLS, where \( E_2 \) and \( E_3 \) are dummy variables indicating membership of the second and third educational classes, respectively. Show that the OLS estimates are:

\[
\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 - \bar{Y}_1 \\ \bar{Y}_3 - \bar{Y}_1 \end{bmatrix}
\]

where \( \bar{Y}_i \) denotes the mean value of \( Y \) in the \( i^{th} \) educational class.
Const \( E_2, E_3 \)  
\[
x = \begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 0 \\
1 & 1 & 0 \\
0 & 0 & 0 \\
\vdots & \vdots & \vdots \\
1 & 0 & 1 \\
1 & 0 & 1 \\
\vdots & \vdots & \vdots 
\end{bmatrix} \Rightarrow x^T x = \begin{bmatrix}
1 & 1 & \ldots & 1 & \ldots & 1 \\
0 & 0 & \ldots & 1 & \ldots & 1 \\
0 & 0 & \ldots & 0 & \ldots & 1 \\
\vdots & \vdots & \ldots & \vdots & \ldots & \vdots 
\end{bmatrix}
\]

\[
\sum_{i=1}^{3} n_i = n
\]

\[
\begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 0 \\
\vdots & \vdots & \vdots 
\end{bmatrix}
\]

Find \((x^T x)^{-1}\):

\[
\begin{bmatrix}
n & n_2 & n_3 \\
n_2 & n_2 & 0 \\
n_3 & 0 & n_3 
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 
\end{bmatrix}
\]

Subtract rows 2 and 3 from row 1; realize \( n - n_2 - n_3 = n_1 \)

\[
\begin{bmatrix}
n_2 & n_2 & 0 & 0 & 0 & 0 \\
n_3 & 0 & n_3 & 0 & 0 & 1 \\
0 & 1 & 1 & -1 & -1 & -1 \\
0 & 0 & 1 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots 
\end{bmatrix}
\]

Divide row 1 by \( n_1 \); row 2 by \( n_2 \); row 3 by \( n_3 \)

\[
\begin{bmatrix}
1 & 0 & 0 & \frac{1}{n_1} & -1/n_1 & -1/n_1 \\
1 & 0 & 0 & \frac{1}{n_2} & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & \frac{1}{n_3} 
\end{bmatrix}
\]

Subtract row 1 from row 2 and row 3

\[
\begin{bmatrix}
1 & 0 & 0 & \frac{1}{n_1} & -1/n_1 & -1/n_1 \\
0 & 1 & 0 & -1/n_1 & 1/n_2 + 1/n_1 & 1/n_1 \\
0 & 0 & 1 & -1/n_1 & 1/n_1 & 1/n_3 + 1/n_1 
\end{bmatrix}
\]

\[
\therefore (x^T x)^{-1} = \begin{bmatrix}
\frac{1}{n_1} & -1/n_1 & -1/n_1 \\
-1/n_1 & 1/n_2 + 1/n_1 & 1/n_1 \\
-1/n_1 & 1/n_1 & 1/n_3 + 1/n_1 
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 0 \\
\vdots & \vdots & \vdots 
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 1 \\
1 & 0 & 1 \\
\vdots & \vdots & \vdots 
\end{bmatrix}
\]

\[
\begin{bmatrix}
n & n_2 & n_3 \\
n_2 & n_2 & 0 \\
n_3 & 0 & n_3 
\end{bmatrix}
\]
3. Suppose that $C$ and $Y$ represent per capita consumption and per capita disposable income respectively, and that J.M. Keynes thinks that they are related by the equation of the form:

$$C = bY + u \quad u \sim N(0, \sigma^2)$$

Keynes wants to estimate the parameter $b$ (which measures the marginal propensity to consume). The random variable $u$ is unobservable, but Keynes can obtain $n$ pairs of observations $(C_1, Y_1)$, $(C_2, Y_2)$, ..., $(C_n, Y_n)$, which relate observations of $(C_i, Y_i)$ over $n$ different years.
(i) Find an expression for the ordinary least squares estimator, $\hat{b}_{OLS}$.

(ii) Use ordinary least squares and the following data to estimate the marginal propensity to consume.

<table>
<thead>
<tr>
<th>Year</th>
<th>1930</th>
<th>1931</th>
<th>1932</th>
<th>1933</th>
<th>1934</th>
<th>1935</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_i$</td>
<td>1059</td>
<td>1016</td>
<td>919</td>
<td>897</td>
<td>934</td>
<td>985</td>
</tr>
<tr>
<td>$Y_i$</td>
<td>1128</td>
<td>1077</td>
<td>921</td>
<td>893</td>
<td>952</td>
<td>1035</td>
</tr>
</tbody>
</table>

(i) $\mathbf{x} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}$

$$\mathbf{x}^T \mathbf{x} = \begin{bmatrix} Y_1 & \cdots & Y_n \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \sum_{i=1}^{n} Y_i^2$$

$$\mathbf{x}^T \mathbf{y} = \begin{bmatrix} C_1 & \cdots & C_n \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = \sum_{i=1}^{n} C_i Y_i$$

$$\hat{b}_{OLS} = (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{y} = \frac{\sum_{i=1}^{n} C_i Y_i}{\sum_{i=1}^{n} Y_i^2}$$

(ii) $\hat{b}_{OLS} = 0.9652$
4. The hardness, $y$, of the shells of eggs laid by a certain breed of chickens was assumed to be roughly linearly related to the amount, $x$, of a certain food supplement put into the diet of chickens. The model assumed is the classical linear regression model. Data were collected and are given below:

<table>
<thead>
<tr>
<th>$y_i$</th>
<th>0.70</th>
<th>0.98</th>
<th>1.16</th>
<th>1.75</th>
<th>0.76</th>
<th>0.82</th>
<th>0.95</th>
<th>1.24</th>
<th>1.75</th>
<th>1.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_i$</td>
<td>0.12</td>
<td>0.21</td>
<td>0.34</td>
<td>0.61</td>
<td>0.13</td>
<td>0.17</td>
<td>0.21</td>
<td>0.34</td>
<td>0.62</td>
<td>0.71</td>
</tr>
</tbody>
</table>

(i) Test the hypothesis that $\beta_1 = 1.00$ versus the hypothesis that $\beta_1 \neq 1.00$. Use a Type I error probability of 5 percent.

(ii) Test the hypothesis that $\beta_1 > 1$ versus the hypothesis that $\beta_1 \leq 1$.

(i) $H_0: \beta_1 = 1.00$
$H_a: \beta_1 \neq 1.00$

$$R = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

$$q = \begin{bmatrix} 1.00 \end{bmatrix}$$

$$F = \frac{\left( R\hat{\beta} - q \right)^T \left[ R(x^T x)^{-1} R^T \right]^{-1} (R\hat{\beta} - q)}{\hat{\varepsilon}^T \hat{\varepsilon} / (n - k)} \sim F_{m,n-k}$$

Rejection region: $F < \text{FINV}(0.05, 1, 8) = 7.5709$ or $F > \text{FINV}(0.95, 1, 8) = 0.0010$

$\therefore F = 415.7 > 7.5709 \Rightarrow$ reject $H_0$ and conclude $\beta_1 \neq 1.00$

(ii) $H_0: \beta_1 > 1.00$
$H_a: \beta_1 \leq 1.00$

Same set up, but now rejection region for $\beta_1 = 1.00$: $F > \text{FINV}(0.05, 1, 8) = 5.318$

$F = 415.7 > 5.318 \therefore$ don't reject $H_0$ (not sufficient evidence to say $\beta_1 \leq 1.00$)
5. A production function model is specified as:
\[ Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \]
where \( Y_i \) = log output, \( X_{2i} \) = log labor input, and \( X_{3i} \) = log capital input. The data refer to a sample of 23 firms, and the observations are measured as from the sample means:
\[
\begin{align*}
\sum x_{2i}^2 &= 12 & \sum x_{2i} x_{3i} &= 8 \\
\sum x_{3i}^2 &= 12 & \sum y_i x_{2i} &= 10 \\
\sum y_i^2 &= 10 & \sum y_i x_{3i} &= 8 \\
\end{align*}
\]

(i) Estimate \( \beta_2, \beta_3 \), their standard errors, and \( R^2 \).
(ii) Test the hypothesis that \( \beta_2 + \beta_3 = 1 \).
(iii) Suppose now that you want to impose the restriction that \( \beta_2 + \beta_3 = 1 \). What is the least squares estimate of \( \beta_2 \) and its standard error? What is the value of \( R^2 \) in this case? Compare the results with those obtained in part (i) and comment.

(i) "as from the sample means" \( \Rightarrow \sum x_{2i} = \sum x_{3i} = \sum y_i = 0 \)
\[
\hat{\beta} = (x^T x)^{-1} x^T y
\]
\[
x^T x = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
1 & x_{21} & \cdots & x_{2n} \\
1 & x_{31} & \cdots & x_{3n} \\
1 & x_{2n} & \cdots & x_{3n}
\end{bmatrix}
\quad \begin{bmatrix}
x_{21} & x_{31} \\
x_{22} & x_{32} \\
x_{2n} & x_{3n}
\end{bmatrix} = \begin{bmatrix}
n \\
\sum x_{2i} \\
\sum x_{3i}
\end{bmatrix} = \begin{bmatrix}
\sum x_{2i}^2 \\
\sum y_i x_{2i} \\
\sum y_i x_{3i}
\end{bmatrix}
\]
\[
\begin{bmatrix}
23 & 0 & 0 \\
0 & 12 & 8 \\
0 & 8 & 12
\end{bmatrix}
\quad \begin{bmatrix}
0.0435 & 0 & 0 \\
0 & 0.15 & -0.1 \\
0 & -0.1 & 0.15
\end{bmatrix}
\]
\[
(x^T x)^{-1} = \begin{bmatrix}
0 & 0.15 & -0.1 \\
0 & -0.1 & 0.15
\end{bmatrix}
\]
\[
x^T y = \begin{bmatrix}
\sum y_i \\
\sum y_i x_{2i} \\
\sum y_i x_{3i}
\end{bmatrix} = \begin{bmatrix}
0 \\
10 \\
8
\end{bmatrix}
\]
\[
(x^T x)^{-1} x^T y = \begin{bmatrix}
0.0435 & 0 & 0 \\
0 & 0.15 & -0.1 \\
0 & -0.1 & 0.15
\end{bmatrix} \begin{bmatrix}
0 \\
10 \\
8
\end{bmatrix} = \begin{bmatrix}
0 \\
0.7 \\
0.2
\end{bmatrix} \Rightarrow \hat{\beta}_{OLS} = \begin{bmatrix}
\hat{\beta}_1 \\
\hat{\beta}_2 \\
\hat{\beta}_3
\end{bmatrix} = \begin{bmatrix}
0 \\
0.7 \\
0.2
\end{bmatrix}
\]
\( \hat{\beta}_{OLS} \sim N(\beta, \sigma^2(x^T x)^{-1}) \)

\( \hat{\sigma}^2 = \frac{1}{n} \hat{\varepsilon}^T \hat{\varepsilon} = 1.4/23 = 0.06087 \)

Note 1: \( \hat{\varepsilon}^T \hat{\varepsilon} = \sum_{i=1}^{n} e_i^2 = \text{RSS} \) (see below)

Note 2: This \( \hat{\sigma}^2 \) is based on MLE for \( \sigma^2 \), but could use \( s^2 = \frac{1}{n-k} \hat{\varepsilon}^T \hat{\varepsilon} \) for unbiased estimator in which case \( s^2 = 1.4/20 = 0.07 \)

\[ \hat{\sigma}^2 (x^T x)^{-1} = \frac{1.4}{23} \begin{bmatrix} 0.0435 & 0 & 0 \\ 0 & 0.15 & -0.1 \\ 0 & -0.1 & 0.15 \end{bmatrix} \begin{bmatrix} 0.002647 & 0 & 0 \\ 0 & 0.00913 & -0.00609 \\ 0 & -0.00609 & 0.00913 \end{bmatrix} \]

Using \( s^2 (x^T x)^{-1} \), \( \frac{1.4}{20} \begin{bmatrix} 0.0435 & 0 & 0 \\ 0 & 0.15 & -0.1 \\ 0 & -0.1 & 0.15 \end{bmatrix} = \begin{bmatrix} 0.003043 & 0 & 0 \\ 0 & 0.0105 & -0.007 \\ 0 & -0.007 & 0.0105 \end{bmatrix} \)

\[ \hat{\sigma}_{\beta_1} = \hat{\sigma}_{\beta_2} = \text{sqrt}(0.00913) = 0.0956 \]

\[ s_{\beta_1} = s_{\beta_2} = \text{sqrt}(0.0105) = 0.1025 \]

\[
\text{RSS} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_{2i} - \hat{\beta}_3 x_{3i})^2 = \sum_{i=1}^{n} (y_i - 0.7 x_{2i} - 0.2 x_{3i})^2 = \\
\sum_{i=1}^{n} (y_i^2 - 0.7 y_i x_{2i} - 0.2 y_i x_{3i} - 0.7 y_i x_{2i} + 0.49 x_{2i}^2 + 0.14 x_{2i} x_{3i} - 0.2 y_i x_{3i} + 0.14 x_{2i} x_{3i} + 0.04 x_{3i}^2) = \\
= \sum_{i=1}^{n} (y_i^2 - 1.4 y_i x_{2i} - 0.4 y_i x_{3i} + 0.49 x_{2i}^2 + 0.28 x_{2i} x_{3i} + 0.04 x_{3i}^2) = \\
= \sum_{i=1}^{n} y_i^2 - 1.4 \sum_{i=1}^{n} y_i x_{2i} - 0.4 \sum_{i=1}^{n} y_i x_{3i} + 0.49 \sum_{i=1}^{n} x_{2i}^2 + 0.28 \sum_{i=1}^{n} x_{2i} x_{3i} + 0.04 \sum_{i=1}^{n} x_{3i}^2 = \\
10 - 1.4(10) - 0.4(8) + 0.49(12) + 0.28(8) + 0.04(12) = 1.4
\]

\[
\text{TSS} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 = 10
\]

\[ R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - 1.4/10 = \boxed{R^2 = 0.86} \]

(ii) \( H_0: \beta_2 + \beta_3 = 1 \)

\( H_a: \beta_2 + \beta_3 \neq 1 \)

\[ \text{R} = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \]

\[ q = \begin{bmatrix} 1 \end{bmatrix} \]

\[
F = \frac{(\hat{R}\hat{\beta} - q)^T [R(x^T x)^{-1} R^T](\hat{R}\hat{\beta} - q)}{\hat{\varepsilon}^T \hat{\varepsilon}/(n-k)} \sim F_{m,n-k}
\]

Rejection region: \( F < \text{FINV}(0.975,1,20) = 0.0010 \) or \( F > \text{FINV}(0.025,1,20) = 5.87 \)
\[
F = \frac{-0.1(1/0.1)(-0.1)}{0.07} = 1.429 \quad \therefore \text{do not reject } H_0
\]

(iii) \( y_i = \beta_1 + \beta_2 x_{2i} + (1 - \beta_2) x_{3i} \Rightarrow y_i - x_{3i} = \beta_1 + \beta_2 (x_{2i} - x_{3i}) \)

\[
\hat{y} = \begin{bmatrix} y_1 - x_{31} \\ \vdots \\ y_n - x_{3n} \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} 1 & x_{21} - x_{31} \\ \vdots \\ 1 & x_{2n} - x_{3n} \end{bmatrix}
\]

\[
\hat{\beta} = \left( \bar{x}^T \bar{x} \right)^{-1} \bar{x}^T \hat{y} = \begin{bmatrix} \sum_{i=1}^{n} y_i - x_{3i} \\ \vdots \\ \sum_{i=1}^{n} y_n - x_{3n} \end{bmatrix} = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^{n} x_{2i} - x_{3i} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^{n} x_{2n} - x_{3n} \end{bmatrix}
\]

\[
\hat{\beta}_1 = \frac{1}{n} \sum_{i=1}^{n} x_{2i} - x_{3i} = \frac{23}{23} = 1 \quad \hat{\beta}_2 = \frac{1}{n} \sum_{i=1}^{n} x_{2i} - x_{3i} = 0
\]

\[
\hat{\beta}_OLS = \left[ \begin{array}{c} \hat{\beta}_1 \\ \hat{\beta}_2 \end{array} \right] = \left[ \begin{array}{c} 1 \\ 0 \end{array} \right]
\]

\[
\hat{\sigma}^2 = \frac{1}{n} \hat{\beta}^T \hat{\epsilon} = 1.5/23
\]

Note 1: \( \hat{\epsilon}^T \hat{\epsilon} = \sum_{i=1}^{n} e_i^2 = \text{RSS} \) (see below)
Note 2: This $\hat{\sigma}^2$ is based on MLE for $\sigma^2$, but could use $s^2 = \frac{1}{n-k}\hat{e}^T\hat{e}$ for unbiased estimator in which case $s^2 = 1.5/21$

$$\hat{\sigma}^2 (\tilde{x}^T \tilde{x})^{-1} = 1.5\begin{bmatrix} 0.0435 & 0 \\ 0 & 0.125 \end{bmatrix} = \begin{bmatrix} 0.002836 & 0 \\ 0 & 0.0082152 \end{bmatrix}$$

Using $s^2 (\tilde{x}^T \tilde{x})^{-1}$, $\frac{1.5}{21}\begin{bmatrix} 0.0435 & 0 \\ 0 & 0.125 \end{bmatrix} = \begin{bmatrix} 0.003106 & 0 \\ 0 & 0.008929 \end{bmatrix}$

\[ \therefore \hat{\sigma}_{\hat{\beta}_2} = \text{sqrt}(0.008152) = 0.0903 \]
\[ s_{\hat{\beta}_2} = \text{sqrt}(0.008929) = 0.0945 \]

\[ \text{RSS} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[ (y_i - \hat{x}_i) - \hat{\beta}_1 - \hat{\beta}_2 (x_{2i} - x_{3i}) \right]^2 = \sum_{i=1}^{n} \left[ y_i - 0.75x_{2i} - 0.25x_{3i} \right]^2 = \]

\[ \sum_{i=1}^{n} (y_i^2 - 1.5y_i x_{2i} - 0.5y_i x_{3i} + 0.5625x_{2i}^2 + 0.375x_{2i} x_{3i} + 0.0625x_{3i}^2) = \]

\[ \sum_{i=1}^{n} y_i^2 - 1.5\sum_{i=1}^{n} y_i x_{2i} - 0.5\sum_{i=1}^{n} y_i x_{3i} + 0.5625\sum_{i=1}^{n} x_{2i}^2 + 0.375\sum_{i=1}^{n} x_{2i} x_{3i} + 0.0625\sum_{i=1}^{n} x_{3i}^2 = \]

\[ 10 - 1.5(10) - 0.5(8) + 0.5625(12) + 0.375(8) + 0.0625(12) = 1.5 \]

\[ \text{TSS} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 = 10 \]

\[ R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - 1.5/10 = \boxed{R^2 = 0.85} \]

In order to enforce $\beta_2 + \beta_3 = 1$, we had to make both parameters bigger. As the numbers turn out, the difference is shared equally between them (0.05 larger than in part i). Why that's the case, I have no idea. It does makes sense that forcing this condition results in a lower $R^2$ value (0.86 vs. 0.85) because we are no longer using all the data to minimize the sum of the squared residuals (we're still minimizing them, but the altered residuals aren't the same). One good thing about the second version is that the parameters are not correlated as they were in part i. I know the independent variables ($x_2$ and $x_3$) are not supposed to be correlated, but we didn't talk about the parameters. The strange thing is that the standard errors for $\beta_2$ is actually smaller in the second case (although for $\beta_1$ it's higher).

6. You will be e-mailed a dataset on heights, sec, mother's height, and father's height.
   (i) Using this dataset, estimate the unconstrained regression of height on sex, mother's height, and father's height. Interpret your results.
   (ii) You wish to test the hypothesis that the coefficients on mother's height and father's height sum to one, and that the coefficient on sex is equal to -8. What is the R matrix and the r vector that correspond to these restrictions?
   (iii) Conduct a test of this hypothesis that has an F-statistic. What are your conclusions?
(i) Height = 55.28 - 12.92 Gender + 0.317 Mom + 0.406 Dad + ε

On initial inspection it would appear gender is the biggest determinant of height, but it's not. What gender does imply is that women (gender = 1) are almost 13 cm shorter than men (assuming their parent's are the same height). This 13 cm isn't much when we factor in the average heights of moms and dads which have a greater impact on height than gender does (in absolute terms): + 51.8 for average mom and + 72.4 for average dad. Father's height is more significant because father's are taller in general, but also more significant on the margin (for each cm of height a father adds 0.406 cm to his offspring versus only 0.317 cm a mom passes on).

(ii) $H_0$: Mom + Dad = 1 and Gender = -8  
$H_a$: one (or both) of these don't hold

\[
R = \begin{bmatrix}
0 & 1 & 1 \\
1 & 0 & 0
\end{bmatrix}
\]

\[
q = \begin{bmatrix}
1 \\
-8
\end{bmatrix}
\]

(iii) Using Stata

```
. test mom + dad = 1
   ( 1)  mom + dad = 1
          F(  1,    29) =  2.95
          Prob > F = 0.0965

. test gender = -8, accumulate
   ( 1)  mom + dad = 1
   ( 2)  gender = -8
          F(  2,    29) =  7.44
          Prob > F = 0.0025
```

Up to a 99.75% confidence level, we can reject that both Mom + Dad = 1 and Gender = -8. Based on the results of the first test, it would appear Gender = -8 is the part that causes the joint test to fail.

Documentation

Prof Werner showed how to find $x^Tx$ for problem 2 in class.

Prof Werner told me "as from the sample means" $\sum x_{2i} = \sum x_{3i} = \sum y_i = 0$ in problem 5.

She also told me how to set it up to do part iii.

Scott showed us how to use the `insheet, regress and test` commands in Stata to do problem 6.