1. Derive the formula for the confidence interval of the individual parameter $\beta_0$ from the formula for the confidence interval of a linear combination of parameters:

$$t^T b \pm t_{\alpha/2}s\sqrt{t^T(X^TX)^{-1}t}.$$ 

**Solution:** We set

$$t = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

so that $t^T b = b_0$. Then $t^T(X^TX)^{-1}t$ is equal to the top left element of $(X^TX)^{-1}$. But this is $c_{00}$ in our notation, so the confidence interval for $\beta_0$ is

$$b_0 \pm t_{\alpha/2}s\sqrt{c_{00}}.$$

2. We model the energy consumption of a household in terms of the household income. The data we collect is:

<table>
<thead>
<tr>
<th>Income ($k$)</th>
<th>Energy consumption ($\times 10$ Btu/yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>1.8</td>
</tr>
<tr>
<td>30</td>
<td>3.0</td>
</tr>
<tr>
<td>40</td>
<td>4.8</td>
</tr>
<tr>
<td>55</td>
<td>5.0</td>
</tr>
<tr>
<td>60</td>
<td>6.5</td>
</tr>
</tbody>
</table>

- Find a 95% confidence interval for the average energy consumption of households with yearly income $50,000$. You may use $t_{0.025} = 3.182$ for 3 degrees of freedom.

**Solution:**

$$X^T X = \begin{bmatrix} 5 & 205 \\ 205 & 9525 \end{bmatrix}, \quad X^T y = \begin{bmatrix} 21.1 \\ 983.0 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1.701 & -0.036 \\ -0.036 & 0.00089 \end{bmatrix}, \quad b = \begin{bmatrix} -0.096 \\ 0.105 \end{bmatrix}$$
The confidence interval is

\[
[1.50 - 0.096, 0.105] \pm 3.182 \sqrt{0.359} \sqrt{1 + [1 50]} \left[ \begin{array}{cc}
1.701 & -0.036 \\
-0.036 & 0.00089
\end{array} \right] \left[ \begin{array}{c}
1 \\
50
\end{array} \right] = (4.17, 6.16).
\]

- Find a 95% prediction interval for the energy consumption of a randomly selected household with yearly income $50,000.

**Solution:**

The prediction interval is

\[
[1.50 - 0.096, 0.105] \pm 3.182 \sqrt{0.359} \sqrt{1 + [1 50]} \left[ \begin{array}{cc}
1.701 & -0.036 \\
-0.036 & 0.00089
\end{array} \right] \left[ \begin{array}{c}
1 \\
50
\end{array} \right] = (3.02, 7.32).
\]

3. Prove that

\[
\frac{(\mathbf{b} - \beta)^T X^T X (\mathbf{b} - \beta)}{\sigma^2}
\]

has a \(\chi^2\) distribution with \(p\) degrees of freedom. (*Hint: you will need a corollary on the distribution of a quadratic form*).

**Solution:** \(\mathbf{b} - \beta\) is a \(p \times 1\) normal random vector with mean \(\mathbf{0}\) and variance \(V = (X^T X)^{-1}\sigma^2\). Therefore \((\mathbf{b} - \beta)^T V^{-1} (\mathbf{b} - \beta)\) has a \(\chi^2\) distribution with \(p\) degrees of freedom and noncentrality parameter \(\frac{1}{2} \mathbf{0}^T V^{-1} \mathbf{0} = 0\). But \(V^{-1} = \frac{1}{\sigma^2} X^T X\), so the result follows.

4. Using the data from question 2, find a joint 95% confidence region for the two parameters \(\beta_0\) and \(\beta_1\). You may keep your answer as an implicit inequality, and use \(f_{0.05} = 9.55\) for 2 and 3 degrees of freedom.

**Solution:**

\[
\frac{(\mathbf{b} - \beta)^T X^T X (\mathbf{b} - \beta)}{\sigma^2} \leq p s^2 f_0
\]

\[
\begin{bmatrix}
-0.096 - \beta_0 \\
0.105 - \beta_1
\end{bmatrix}
\begin{bmatrix}
5 & 205 \\
205 & 9525
\end{bmatrix}
\begin{bmatrix}
-0.096 - \beta_0 \\
0.105 - \beta_1
\end{bmatrix}
\leq 2 \times 3.182 \times 0.359 \times 9.55
\]

\[
5 \beta_0^2 + 9525 \beta_1^2 + 410 \beta_0 \beta_1 - 42.1 \beta_0 - 1960.9 \beta_1 + 100.9 \leq 6.86.
\]

5. We can write \(SS_{\text{Reg}}\) and \(SS_{\text{Res}}\) as

\[
SS_{\text{Reg}} = y^T X (X^T X)^{-1} X^T y
\]

and

\[
SS_{\text{Res}} = y^T [I - X (X^T X)^{-1} X^T] y.
\]
Show that these two quadratic forms are independent.

**Solution:** From the theorem for independence of quadratic forms,

\[
AVB = X(X^T X)^{-1}X^T \sigma^2 I [I - X(X^T X)^{-1}X^T] \\
= \sigma^2 (X(X^T X)^{-1}X^T - X(X^T X)^{-1}X^T X(X^T X)^{-1}X^T) \\
= \sigma^2 (X(X^T X)^{-1}X^T - X(X^T X)^{-1}X^T) \\
= 0.
\]

Therefore the quadratic forms are independent.

6. Using the data from question 2, test the null hypothesis \( H_0 : \beta = 0 \).

**Solution:**

\[
SS_{Res} = 1.08, SS_{Reg} = 101.5.
\]

Our test statistic is

\[
\frac{SS_{Reg}/p}{SS_{Res}/(n-p)} = 141.3 > 9.55.
\]

Therefore, we reject the model hypothesis — the model is adequate.