Problem 1  (25 points)
Consider the multiple regression model

\[ y = X\beta + \epsilon, \]

with \( E(\epsilon) = 0 \) and \( \text{cov}(\epsilon) = \sigma^2 I. \)

Answer the following questions:

a. Show that the regression sum of squares can be expressed as:

\[ SSR = \hat{\beta}'X'X\hat{\beta} - ny^2. \]

*Hint:* Express \( SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \) in vector notation.

b. Find \( E(\epsilon'\epsilon). \) Note: Do not assume that \( \epsilon \) is normal (so you cannot use the \( \chi^2 \) distribution)!
Problem 2  (25 points)

After fitting the multiple regression model \( y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i \), where \( E(\epsilon_i) = 0 \), \( E(\epsilon_i \epsilon_j) = 0 \) for \( i \neq j \), and \( \text{var}(\epsilon_i) = \sigma^2 \), to a data set with \( n = 15 \) observations it is found that \( s^2_e = 3 \) and

\[
(X'X)^{-1} = \begin{pmatrix}
0.5 & 0.3 & 0.2 & 0.6 \\
0.3 & 6.0 & 0.5 & 0.4 \\
0.2 & 0.5 & 0.2 & 0.7 \\
0.6 & 0.4 & 0.7 & 3.0
\end{pmatrix}.
\]

Answer the following questions:

a. Find the estimate of \( \text{var}(\hat{\beta}_1) \).

b. Find the estimate of \( \text{cov}(\hat{\beta}_1, \hat{\beta}_3) \).

c. Find the estimate of \( \text{corr}(\hat{\beta}_1, \hat{\beta}_3) \).

d. Find the estimate of \( \text{var}(\hat{\beta}_1 - \hat{\beta}_3) \).

e. Give the procedure that will test the hypothesis:

\[ H_0 : \beta_0 - 2\beta_1 + 3\beta_2 - 5\beta_3 = 0 \]
\[ H_a : \beta_0 - 2\beta_1 + 3\beta_2 - 5\beta_3 \neq 0 \]

Please show all your work! No calculations are necessary.
Problem 3  (25 points)
Part A:
Consider the simple regression model through the origin: \( y_i = \beta_1 x_i + \epsilon_i \), with \( E(\epsilon_i) = 0, \) \( \text{var}(\epsilon_i) = \sigma^2, \) \( E\epsilon_i \epsilon_j = 0, \) \( \text{and} \) \( \epsilon_i \sim N(0, \sigma). \) Let \( \hat{\beta}_1 \) and \( \hat{\sigma}^2 \) be the maximum likelihood estimators of \( \beta_1 \) and \( \sigma^2 \) respectively. Answer the following questions:

a. Find the Fisher information matrix for the vector \( \theta = (\beta_1, \sigma^2)' \).

b. What is the asymptotic distribution of \( \hat{\theta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\sigma}^2 \end{pmatrix} \). Write the pdf of this distribution.

Part B:
The multiple regression model can be expressed as \( y_i = x_i' \beta + \epsilon_i \) with \( i = 1, \ldots, n \) where \( E(\epsilon_i) = 0, \) \( \text{var}(\epsilon_i) = \sigma^2, \) \( \text{and} \) \( \text{cov}(\epsilon_i, \epsilon_j) = 0 \) when \( i \neq j \). Note: \( x_i \) is the \( i \)th row of matrix \( X \). Answer the following questions:

a. Show that \( \hat{y}_i = x_i' \hat{\beta} \) is unbiased estimator of \( x_i' \beta \).

b. What is the variance of \( \hat{y}_i \)? Please remember this is a multiple regression model.

c. Does there exist any other linear unbiased estimator of \( x_i' \beta \) (say \( \hat{y}_i = x_i' b \)), with smaller variance than the variance of the estimator \( \hat{y}_i \)?
Problem 4 (25 points)
Researchers 1 and 2 were working independently on similar problems.

Using \( n_1 \) data points, researcher 1 formed the model \( y_1 = X_1 \beta + \epsilon_1 \), where \( y_1 \) is \( n_1 \times 1 \), \( X_1 \) is \( n_1 \times (k+1) \), \( \beta \) is \( (k+1) \times 1 \), and \( \epsilon_1 \) is \( n_1 \times 1 \), with \( E(\epsilon_1) = 0 \) and \( \text{cov}(\epsilon_1) = \sigma^2 I \).

Using \( n_2 \) data points, researcher 2 formed the model \( y_2 = X_2 \beta + \epsilon_2 \), where \( y_2 \) is \( n_2 \times 1 \), \( X_2 \) is \( n_2 \times (k+1) \), \( \beta \) is \( (k+1) \times 1 \), and \( \epsilon_2 \) is \( n_2 \times 1 \), with \( E(\epsilon_2) = 0 \) and \( \text{cov}(\epsilon_2) = \sigma^2 I \).

Note: Each researcher is trying to estimate the same coefficient vector \( \beta \).

Answer the following questions:

a. Suppose that the researchers worked independently. Give \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \), their separate least squares estimates.

b. Suppose that they cooperate and pool their data. So now the model will be:

\[
\begin{pmatrix}
  y_1 \\
  y_2 \\
\end{pmatrix} =
\begin{pmatrix}
  X_1 \\
  X_2 \\
\end{pmatrix} \beta +
\begin{pmatrix}
  \epsilon_1 \\
  \epsilon_2 \\
\end{pmatrix}.
\]

Show that their combined least squares estimates are:

\[ \hat{\beta} = (X_1'X_1 + X_2'X_2)^{-1}(X_1'y_1 + X_2'y_2). \]

c. Consider the case \( n_1 = n_2 \) and \( X_1 = X_2 \), but \( y_1 \) not necessarily equals to \( y_2 \). Now answer question (b) again.

d. Consider the situation in (c). Find the variance covariance matrix of \( \hat{\beta} \).