For the simple linear regression model, show that the off-diagonal elements of the hat matrix are
\[ h_{ij} = \frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{S_{XX}} \]
where \( S_{XX} = \sum_{i=1}^{n} (x_i - \bar{x})^2 \).

Consider a fixed effect multiple linear regression model \( y = X\beta + \epsilon \). Suppose that the probability distribution of the random error term is as follows:
\[ f(\epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\epsilon/\sigma)^2}; -\infty < \epsilon < \infty \]
Using the method of maximum likelihood estimation, derive the estimators for \( \beta \) and \( \sigma^2 \).

For the model \( y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon \), write the \( T \) matrix for testing the following set of hypotheses:

\[ H_0: \beta_1 = \beta_4 = \beta_5 \text{ and } \beta_2 = \beta_3 = \beta_6 \]
\[ H_1: \text{ At least one of the conditions in } H_0 \text{ is not true} \]

Consider the following regression model:
\[ y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon \]
A sample of size \( n = 35 \) gave \( R^2 = 0.73 \). The investigator knows that the standard deviation of \( y \) is 4.4. Find the adjusted \( R^2 \).

<table>
<thead>
<tr>
<th>Q. No.</th>
<th>Question</th>
<th>Full Marks</th>
<th>Marks Obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>For the simple linear regression model, show that the off-diagonal elements of the hat matrix are ( h_{ij} = \frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{S_{XX}} ) where ( S_{XX} = \sum_{i=1}^{n} (x_i - \bar{x})^2 ).</td>
<td>08</td>
<td></td>
</tr>
</tbody>
</table>
| 2      | Consider a fixed effect multiple linear regression model \( y = X\beta + \epsilon \). Suppose that the probability distribution of the random error term is as follows: \[ f(\epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\epsilon/\sigma)^2}; -\infty < \epsilon < \infty \]
Using the method of maximum likelihood estimation, derive the estimators for \( \beta \) and \( \sigma^2 \). | 08 |  |
| 3      | For the model \( y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon \), write the \( T \) matrix for testing the following set of hypotheses:
\[ H_0: \beta_1 = \beta_4 = \beta_5 \text{ and } \beta_2 = \beta_3 = \beta_6 \]
\[ H_1: \text{ At least one of the conditions in } H_0 \text{ is not true} \] | 04 |  |
| 4      | Consider the following regression model:
\[ y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon \]
A sample of size \( n = 35 \) gave \( R^2 = 0.73 \). The investigator knows that the standard deviation of \( y \) is 4.4. Find the adjusted \( R^2 \). | 05 |  |

**Instructions:**

1. Use of mobiles is not allowed in exam. If you have your mobile with you, turn it off and place it on the table so that it is visible to proctor.

2. Show all the calculation steps. There are points for each step.
## STAT 310: Linear Regression

Term 191, Second Major Exam (Computational)
Sunday November 03, 2019 (7:00 pm)

Name: ______________________  ID #: ______________________

<table>
<thead>
<tr>
<th>Question No</th>
<th>Full Marks</th>
<th>Marks Obtained</th>
<th>Question No</th>
<th>Full Marks</th>
<th>Marks Obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>5a</td>
<td>07</td>
<td></td>
<td>6d</td>
<td>03</td>
<td></td>
</tr>
<tr>
<td>5b</td>
<td>05</td>
<td></td>
<td>6e</td>
<td>05</td>
<td></td>
</tr>
<tr>
<td>6a</td>
<td>03</td>
<td></td>
<td>6f</td>
<td>04</td>
<td></td>
</tr>
<tr>
<td>6b</td>
<td>04</td>
<td></td>
<td>6g</td>
<td>04</td>
<td></td>
</tr>
<tr>
<td>6c</td>
<td>04</td>
<td></td>
<td>6h</td>
<td>04</td>
<td></td>
</tr>
<tr>
<td><strong>Total (out of 43)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Q.No.5: - (7+5 = 12 points) Consider the data on 40 paired observations \((x, y)\). The functional form of the relationship between the 2 variables is given as \(y = \frac{e^x}{\delta + ye^x}\).

(a) Estimate the unknown parameters \(\delta\) and \(\gamma\) and predict the value of \(y\) when \(x = 11.3\).

(b) Check for the presence of any outlier(s) using externally studentized residuals \((t_i)\). Write down your conclusions.
An engineer studied the effect of four variables on a dimensionless factor used to describe pressure drops in a screen-plate bubble column. Table B.9 summarizes the experimental results.

a) Fit a multiple linear regression model relating this dimensionless factor to these regressors. Write down the estimated model.

b) Test for significance of regression. What conclusions can you draw?

c) Use t tests to assess the contribution of each regressor to the model. Discuss your findings.
d) Calculate $R^2$ and $R^2_{adj}$ for this model. Discuss your results.

e) Test the hypothesis that $x1$ and $x4$ are insignificant against the alternative that at least one of them is significant.
f) Find a 99% confidence interval for the regression coefficient of $x_2$.

g) Construct a 99% confidence interval for the average dimensionless factor when $x_1=5.6$, $x_2=10.1$, $x_3=0.34$ and $x_4=0.59$.

h) Is there multicollinearity in the model? Write down your criteria and make conclusions.