Exam I

You have one hour for this exam. The exam is worth 12.5 percent of your grade. The exam is closed book, but calculators are permitted. **Only the calculators that I pass out are to be used.** Please do all your work carefully and check your calculations before handing your work in. **PLEASE SHOW ALL YOUR WORK TO AID IN THE AWARDING OF PARTIAL CREDIT WHERE APPLICABLE.**

\[ y = b_1 + b_2 x \]

\[ \bar{y} = b_1 + b_2 \bar{x} \]

\[ \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \]

\[ \sum e_i^2 \]

\[ \sum (y_i - \bar{y})^2 = 1 \]

\[ \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2} \]

\[ s_{\hat{b}_2} \]

\[ s_{\hat{b}_2} = \sqrt{\frac{s^2}{b_2}} \]

\[ s_{\hat{b}_2}^2 = \frac{s^2}{\sum (x_i - \bar{x})^2} \]

\[ s^2 = \frac{\sum e_i^2}{n-2} = \frac{\sum (y_i - \bar{y})^2 - b_2 \sum (x_i - \bar{x})(y_i - \bar{y})}{n-2} \]
(45 points.) Suppose that I did a survey of five students from my ECON 110 class of last semester. I found out their Math SAT score and their score on my final exam. The results are given below.

<table>
<thead>
<tr>
<th>M-SAT</th>
<th>Final</th>
<th>( (x_i - \bar{x}) )</th>
<th>( (y_i - \bar{y}) )</th>
<th>( (x_i - \bar{x})^2 )</th>
<th>( (x_i - \bar{x})(y_i - \bar{y}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>75</td>
<td>0</td>
<td>-12</td>
<td>(-12)</td>
<td>120</td>
</tr>
<tr>
<td>700</td>
<td>95</td>
<td>10</td>
<td>8</td>
<td>(100)</td>
<td>80</td>
</tr>
<tr>
<td>600</td>
<td>90</td>
<td>0</td>
<td>3</td>
<td>(9)</td>
<td>(27)</td>
</tr>
<tr>
<td>550</td>
<td>85</td>
<td>5</td>
<td>2</td>
<td>(25)</td>
<td>(105)</td>
</tr>
<tr>
<td>650</td>
<td>90</td>
<td>15</td>
<td>3</td>
<td>(225)</td>
<td>(322)</td>
</tr>
</tbody>
</table>

\[ \bar{x} = \frac{\sum x_i}{n} = \frac{500 + 700 + 600 + 550 + 650}{5} = 600 \]

\[ \bar{y} = \frac{\sum y_i}{n} = \frac{75 + 95 + 90 + 85 + 90}{5} = 87 \]

Where:
- **M-SAT** = Score on the Math Portion of the SAT (points)
- **Final** = Final Exam Score (measured as percent)

a. Please run a regression to find the effect that one's Math SAT score has on one's score on the final exam.

\[ b_2 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{2250}{25000} = 0.09 \]

\[ b_1 = \bar{y} - b_2 \bar{x} = 87 - 0.09(600) = 87 - 54 = 33 \]

\[ \text{Final} = 33 + 0.09 \text{ SAT} \]

b. In terms of the variables, explain what the slope tells you.

As the SAT math score of a student improves by one point, we predict their score on the final to increase by 0.09 percent point.

c. In terms of the variables, explain what the constant tells you.

If a student had an SAT math score of zero, we would predict their score on the final exam to be 33.
3 (20 points) Suppose we gather weekly data for a hamburger restaurant over the course of the six months, i.e., 26 observations. We ascertain information on their Revenue, the price of hamburgers and their advertising expenditures. The regression results are:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>104.8</td>
<td>6.5</td>
</tr>
<tr>
<td>Price</td>
<td>-6.7</td>
<td>3.2</td>
</tr>
<tr>
<td>Adv</td>
<td>3.0</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Total Sum of Squares = 13,572.9
Sum of Squared Errors = 1805.2

Rev = Total Revenue in $1,000's
Price = Burger Price in dollars
Adv = Advertising Expenditures in $1,000’s

a. In terms of the variables, explain what the slope on advertising tells us?
   When advertising expenditures rise by one thousand dollars, we predict Total Revenue to rise by 3.0 thousand dollars controlling for the price.

b. What would be the predicted impact of a ten cent increase in the price of hamburgers?
   \[ (1) - \frac{(-6.7)}{3.2} \approx -2.1 \] percent in total revenue.

(c) Can we be at least 95% certain that the price of hamburgers has an impact on total revenue? Please perform a formal test.

\[ H_0 : \beta_2 = 0 \]
\[ H_a : \beta_2 \neq 0 \]

\[ t = \frac{-26.3}{2.1} \approx -12.5 \]

\[ t_{25} = -1.717 \]

\[ -12.5 < -1.717 \]

\[ (2.069 < 2.09 \neq 2.019) \]

(d) In terms of the variables, explain what your result in part c tells you?
   The price has a negative effect on total revenue, controlling for advertising.

(e) Calculate the R² and in terms of the variables explain what it tells you.
   \[ R^2 = 0.867 \]

\[ 86.7 \% \] of the variation in total revenue is explained by advertising.
d. Suppose there was a student in the class who had a 575 on the math SAT. What would you predict their final exam score to be?

\[
\text{Final} = 33 + 0.9 \times \text{SAT} = 33 + 0.9 \times 575
\]

\[
= 33 + 517.5
\]

\[
= 847.5
\]

e. Look at the first student's data. Calculate their error (residual.) (Note: No credit will be given for stating the student’s error was in taking a class with me in the first place ;)

\[
\text{Final} = 33 + 0.9 \times \text{SAT}
\]

\[
= 33 + 0.9 \times 575
\]

\[
= 33 + 517.5
\]

\[
= 847.5
\]

\[
\text{Error} = \text{Final} - \text{Actual}
\]

\[
= 847.5 - 78
\]

\[
= 769.5
\]

2. (14 points) Most of the houses in my neighborhood have 3 or 4 bedrooms. Suppose a survey of 20 houses shows that the number of bedrooms and the number of children is the following. (Each cell represents the number of families. So there are four families with no children and a three bedroom house.)

<table>
<thead>
<tr>
<th>Number of Bedrooms</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

a. What is the relative frequency of families?

\[
\frac{12}{20} = 0.6
\]

b. What is the expected value of the number of bedrooms in my neighborhood?

\[
E(X) = \mu = \sum x_i \cdot P(x_i) = 420(3) + 120(4) = 12 + 24 = 3.6
\]

c. Suppose my town places a tax of $30 + $20*(Number of Bedrooms) on each house. What is the expected value of the tax that families in my neighborhood will pay?

\[
E(\text{Tax}) = E(a + bX) = a + bE(X) = 30 + 20(3.6) = 30 + 72 = 102
\]

d. What is the expected value of the number of bedrooms in my neighborhood for families with 2 children?

\[
E(\text{Bedrooms}) = \sum x_i \cdot P(x_i) = \frac{3}{8} \cdot 4 + \frac{5}{8} \cdot 3 = 2.8 + 2.3 = 5.1
\]
4. (5 points.) Suppose we knew that the mean household income of the population of the United States was $58,000 and the standard deviation was $15,000. If we wanted to see whether we could be sure the people in a town had an income higher than the national average, what test would we use?

When the standard deviation is known, use the $z$-test.

5. (6 points)

a. In a simple regression, state one of the assumptions of the Gauss-Markov Theorem.

- $X$ is unrelated to $U$.
- $\text{E}(U_i) = 0$.

b. State an assumption of the Gauss-Markov Theorem for a multiple regression that is not part of the assumptions for a simple regression.

- No exact linear relationship between independent variables.
- $m > k$ or $m$ observations exceed the number of independent variables.

c. If the assumptions of the Gauss-Markov Theorem are met, what occurs?

The OLS estimators from the sample are unbiased, efficient estimators of the population regression parameters (also could say BLUE).