Midterm Exam (Version A)

The exam is closed book and closed notes. You may use your calculators but please show your work step by step. No graphing calculators or cell phones. **You must show your work to receive full credit**

**Problem 1 (8 points)**
You have a sample of three cars. Each weighs 1800 lbs, 2000 lbs, and 2500 lbs and their EPA fuel economy ratings are 30 mpg, 25 mpg, and 23 mpg, respectively

a. What is the average weight and average fuel economy? (2 points)

b. What is the correlation between weight and fuel economy for the three cars? (2 points)

c. Suppose that you accidentally weighed the heaviest car without removing your kids. The new mean of all three cars with kids removed is 2050 lbs. How much did your kids weigh? (2 points)
d. *(Harder question)* Suppose that the covariance of a different sample of three cars with weight measured in lbs and efficiency in miles per gallon was 30. We are shipping the cars for sale in Canada. What is the covariance if we measure weight in kilograms (kgs) and distance is kilometers (km)? (1 mile=1.61 km, 1 lb=0.45 kilograms) (2 points)

**Problem 2 (6 Points)**

You flip a coin three times. The coin is fair, with Pr(Head)=Pr(Tail)=.5.

a. Draw a Venn Diagram for this experiment (1 points)

b. What is the probability that you will get exactly 2 heads? (2 points)

c. What is the probability that you will get no more than 2 heads OR more than 1 tail? (3 points)
Problem 3  (8 Points)
The average number of miles run per week for UCSC students is characterized by a normal distribution with mean 2.5 and standard deviation 1.

a. What is the probability that a randomly selected student runs 3 miles per week? (1 point)

b. What is the probability that a randomly selected student runs between 2 and 4 miles per week? (3 points)

c. (Harder question) In addition to the distribution described above, the number of study hours per week for UCSC students follows a uniform distribution, with a minimum 0 and maximum 8. Assume that miles run and study hours are independent from one another. What is the probability that a randomly selected student runs between 1 and 2 miles per week OR studies between 4 and 5 hours per day? (4 points)
Problem 4 (8 Points)
On Professor Spearot’s recent month long sabbatical, he played four 10 song gigs at the Red Room. He collected data on the number of original songs performed, “Originals”, and tips earned, “Tips”.

<table>
<thead>
<tr>
<th>Originals</th>
<th>Tips</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>$10</td>
</tr>
<tr>
<td>9</td>
<td>$5</td>
</tr>
<tr>
<td>1</td>
<td>$20</td>
</tr>
<tr>
<td>4</td>
<td>$25</td>
</tr>
</tbody>
</table>

He wants to run the following regression:

\[ \text{Tips} = \beta_0 + \beta_1 \text{Originals} + u \]

a. Please compute \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) (2 points).
b. If assumptions 1-4 from the proof shown in class hold so our estimate of $\hat{\beta}_1$ is an unbiased estimate of $\beta_1$, how much money is gained or lost by playing an additional original song? (2 points)

c. If assumptions 1-4 from the proof shown in class hold so our estimate of $\hat{\beta}_1$ is an unbiased estimate of $\beta_1$, how much do I make in tips if all 10 songs are covers (not original)? (2 points)

d. Compute and interpret the R-squared. (2 points)

Extra Credit: The original songs contain more screaming than the covers (non-original). That is, covers are positively correlated with screaming. Red Room patrons do not like screaming, and thus screaming tends to yield fewer tips. What assumption do we violate by estimating $Tips = \beta_0 + \beta_1Originals + u$? Is the bias positive or negative? (2 points)
Extra Credit: What is the most important thing you will learn in this class (I said this in lecture)? (2 Points)

Helpful Formulas

\[
\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2 \\
\hat{\sigma}_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y) \\
\hat{\rho}_{xy} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y} \\
\hat{\beta}_0 = \hat{\mu}_y - \hat{\beta}_1 \hat{\mu}_x \\
\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y)}{\sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2} \\
R^2 = 1 - \frac{SSR}{SST} \\
SSR = \sum_{i=1}^{n} (\hat{\mu}_i)^2 \\
SST = \sum_{i=1}^{n} (y_i - \hat{\mu}_y)^2
\]