AP Statistics: Chapter 5 Review

1. The data below were gathered on a random sample of 5 basking sharks, swimming through the water and filter-feeding, i.e. letting the water bring food into their mouths.

Mean speeds for basking sharks

<table>
<thead>
<tr>
<th>Body Length (meters)</th>
<th>Mean speed (meters/second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0</td>
<td>0.89</td>
</tr>
<tr>
<td>4.5</td>
<td>0.83</td>
</tr>
<tr>
<td>4.0</td>
<td>0.76</td>
</tr>
<tr>
<td>6.5</td>
<td>0.94</td>
</tr>
<tr>
<td>5.5</td>
<td>0.94</td>
</tr>
</tbody>
</table>

a) What is the value of the correlation coefficient for these data?

b) Find the equation of the least squares line describing the relationship between

\[ x = \text{body length} \] and \[ y = \text{mean speed}. \]
c) If these sharks are representative of the population of basking sharks, what would you predict is the mean speed for a filter-feeding basking shark that is 5.0 meters in length? Show any calculations below.

d) The largest basking shark in the sample is measured as 6.5 meters long. Theory predicts a maximum length of about 12.26 meters. Would it be reasonable to use the equation from part (b) above to predict the mean filter-feeding speed for a 12 meter long basking shark? Why or why not?

2. The preservation of objects made of organic material is a constant concern to those caring for items of historical interest. For example, some delicate fabrics are natural silks, made of protein and biodegradable. Many silks in museum collections are in danger of crumbling. It would be of great benefit to be able to assess the delicacy of the fabric before making decisions about displaying it. One possibility is chemical analysis, which might give some evidence about the brittle nature of a fabric. To investigate this possibility, bio-chemical data in the form of a ratio of the amount of certain amino acids in the fibers was acquired from the linings of sixteen 19th and early 20th century Japanese kimonos, as well as the tenacity (breaking stress) of the fabrics.

<table>
<thead>
<tr>
<th>Amino acid ratio</th>
<th>Tenacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.05</td>
<td>1.20</td>
</tr>
<tr>
<td>1.78</td>
<td>1.60</td>
</tr>
<tr>
<td>2.08</td>
<td>1.30</td>
</tr>
<tr>
<td>2.62</td>
<td>0.90</td>
</tr>
<tr>
<td>2.00</td>
<td>1.80</td>
</tr>
<tr>
<td>1.92</td>
<td>1.60</td>
</tr>
<tr>
<td>1.80</td>
<td>1.30</td>
</tr>
</tbody>
</table>
Using the data from the Japanese kimonos, construct the least squares best fit line for tenacity as a function of the amino acid ratio.

a) What is the equation of the least squares best fit line?

b) Graph the least squares best fit line on the scatter plot that appears on the next page.

c) Approximately what proportion of the variability in tenacity is explained by the amino acid ratio?

The theory of fiber strength suggests that the relationship between fiber tenacity and amino acid ratio is logarithmic, i.e.
\[ \hat{T} = \alpha + \beta \log R \]
where \( T \) is the tenacity and \( R \) is the amino acid ratio. Perform the appropriate transformation of variable(s) and fit this logarithmic model to the data.

Tenacity Vs. AminoAcid Ratio

![Graph](image.png)
d) What is the resulting best fit line using this model?

\[ \text{WrestSpectators} = 0.667 + 0.701 \text{Wrestlers} \]

e) For an amino acid ratio of \( R = 1.5 \), what is the predicted tenacity?

\[ \text{Predicted Tenacity} = \frac{1.5}{R} \]

f) Using your results so far, does it appear that the transformed model is no improvement over the linear model, a slight improvement, or a significant improvement? Justify your response with an appropriate statistical argument.

3. The *Des Moines Register* recently reported the ratings of high school sportsmanship as compiled by the Iowa High School Athletic Association. For each school the spectators and participants were rated by referees, where 1 = superior, and 5 = unsatisfactory. A regression analysis of the average scores given to wrestling spectators and wrestlers is shown below.

**Wrestling Spectators vs. Wrestlers**

- **Linear Fit**
  \[ \text{WrestSpectators} = 0.667 + 0.701 \text{Wrestlers} \]

- **Summary of Fit**
  \[
  \begin{align*}
  \text{RSquare} & \quad 0.467 \\
  \text{RSquare Adj} & \quad 0.465 \\
  s & \quad 0.322
  \end{align*}
  \]
a) Interpret the correlation between the ratings of spectators and wrestlers?

b) Interpret the coefficient of determination.

c) Interpret the value of the standard deviation about the least squares line?