Assignment #2, Stat 5511, Spring, 2008

Due Friday, February 15

(1) The given data on fish survival and ammonia concentration was taken from the paper "Effects of Ammonia on Growth and Survival of Rainbow Trout in Intensive Static-Water Culture" (Trans. of the Amer. Fisheries Soc.(1983).)

<table>
<thead>
<tr>
<th>ammonia exposure</th>
<th>percent survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>85</td>
</tr>
<tr>
<td>10</td>
<td>92</td>
</tr>
<tr>
<td>20</td>
<td>85</td>
</tr>
<tr>
<td>20</td>
<td>96</td>
</tr>
<tr>
<td>25</td>
<td>87</td>
</tr>
<tr>
<td>27</td>
<td>80</td>
</tr>
<tr>
<td>27</td>
<td>90</td>
</tr>
<tr>
<td>31</td>
<td>59</td>
</tr>
<tr>
<td>50</td>
<td>62</td>
</tr>
</tbody>
</table>

\[ x = \text{ammonia exposure} \quad \text{and} \quad y = \text{percent survival of fish.} \]

Use SAS to answer the following questions:

(a) Regress percent survival (y) on ammonia exposure (x) (using the simple linear regression model). Write out the regression equation, and interpret the slope.

(b) Test \( H_0: \beta_1 = 0 \) against \( H_1: \beta_1 \neq 0 \). Use \( \alpha = .1 \).

(c) Construct a 90\% confidence interval for \( \beta_1 \). How does this confirm your two-sided test in (b)?

(d) Construct a 90\% confidence interval for the mean percentage of survival with ammonia exposure = 24.

(e) Construct a 90\% confidence interval for the percentage of survival with ammonia exposure=24.

(f) Construct a scatter diagram and draw in the regression line.

(2) (a) Let \((x_1,y_1), (x_2,y_2), \ldots, (x_n,y_n)\) be \(n\) data points. Consider the simple regression model \( y = \beta_0 + \beta_1 x + \varepsilon \). Find the least squares estimate of \( \beta_1 \) for this model (the model with known intercept \( \beta_0 = 0 \)).

(b) Calculate the value of the least squares estimate of \( \beta_1 \) with the following data:

<table>
<thead>
<tr>
<th>(x_i)</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>(y_i)</td>
<td>3</td>
<td>7</td>
<td>12</td>
<td>16</td>
<td>20</td>
</tr>
</tbody>
</table>

(3) Use SAS to do problem 2.3 on p.54.

(4) In the simple linear regression model: \( y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \), prove that:

(a) \( \hat{\beta}_1^2 \sum (x_i - \bar{x})^2 = \sum (\hat{y}_i - \bar{y})^2 \)

(b) \( SS_T = SS_R + SS_{Res} \), where \( SS_T = \sum (y_i - \bar{y})^2 \), \( SS_R = \sum (\hat{y}_i - \bar{y})^2 \) and \( SS_{Res} = \sum (y_i - \hat{y}_i)^2 \).

(5) (Optional) For the simple linear regression model as in problem (4),

(a) prove that \( \hat{\beta} \) and \( \hat{\beta}_1 \) are uncorrelated, i.e. \( Cov(\hat{\beta}, \hat{\beta}_1) = 0 \).

(b) prove that \( (\hat{\beta}_1 - \beta_1)/(\hat{\sigma}/\sqrt{SS_{XX}}) \) has a t-distribution with (\(n-2\)) degree of freedom.