Statistics 200 Honors  
Homework Solutions  
Solution Set 10: Due dates Nov. 2 – Nov. 5

Page 734, exercise 11.4 Refer to the educational data for 78 seventh-grade students given in Table 1.6 (page 33). We view GPA as the response variable. IQ, gender, and self-concept are the explanatory variables.

Note that there is a typo in the online version of the dataset, which is immediately apparent from looking at a boxplot of the IQ values. Namely, observation 43 should be changed from an IQ of 12 to an IQ of 123.

(a) Find the correlation between GPA and each of the explanatory variables. What percent of the total variation in student GPAs can be explained by the straight-line relationship for each of the explanatory variables?

Here is the relevant output from S+ (a different statistics program from Minitab):

<table>
<thead>
<tr>
<th>GPA</th>
<th>IQ</th>
<th>SELFCONCEPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA 1.0000</td>
<td>0.6337</td>
<td>0.3044</td>
</tr>
<tr>
<td>IQ 0.6337</td>
<td>1.0000</td>
<td>0.3235</td>
</tr>
<tr>
<td>SELFCONCEPT 0.3044</td>
<td>0.3235</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Note that S+ gives the results in the form of a square matrix and does not output the p-values by default. We conclude that the correlation between GPA and IQ is 0.634 and the correlation between GPA and Self-concept is 0.304. In each case, the percent of total variation explained by a simple linear regression is the square of the correlation. Thus, regression on IQ explains 40.2% of the variation of GPA, whereas regression on self-concept explains 9.3%. (NOTE: We cannot therefore add these numbers and claim that together the two variables explain 49.5%. The reason is that there is some overlap.)

(b) The importance of IQ in explaining GPA is not surprising. The purpose of the study is to assess the influence of self-concept on GPA. So we will include IQ in the regression model and ask, “How much does self-concept contribute to explaining GPA after the effect of IQ on GPA is taken into account?” Give a model that can be used to answer this question.

The model is

\[ \text{GPA} = \beta_0 + \beta_1 \text{IQ} + \beta_2 \text{Self-concept}. \]  

(1)

The coefficient \( \beta_2 \) is the size of the effect of self-concept after IQ is taken into account. Thus, the key question will be whether \( \beta_2 \) is nonzero.
(c) Run the model and report the fitted regression equation. What percent of the variation in GPA is explained by the explanatory variables in your model?

The regression output from S+ is as follows:

\[
\begin{align*}
\text{Coefficients:} & \\
& \begin{array}{cccc}
\text{Value} & \text{Std. Error} & t \text{ value} & \text{Pr}(>|t|) \\
(\text{Intercept}) & -3.8155 & 1.5629 & -2.4413 & 0.0170 \\
\text{IQ} & 0.0953 & 0.0149 & 6.3922 & 0.0000 \\
\text{SELFCONCEPT} & 0.0159 & 0.0134 & 1.1862 & 0.2393 \\
\end{array}
\end{align*}
\]

Residual standard error: 1.63 on 75 degrees of freedom
Multiple R-Squared: 0.4126

Using the reported value of R-squared, we conclude that 41.3% of the variation in GPA is explained by the explanatory variables in the model.

(d) Translate the question of interest into appropriate null and alternative hypotheses about the model parameters. Give the value of the test statistic and its p-value. Write a short summary of your analysis with an emphasis on your conclusion.

Within the context of model (1), the null hypothesis is

\[ H_0 : \beta_2 = 0. \]

The value of the test statistic and its p-value come directly from the regression output: The \( t \) statistic equals 1.1862 on 1 degree of freedom, which yields a \( p \)-value of 0.2393.

Despite the fact that self-concept by itself appears to be significantly linearly associated with GPA (according to output not shown here), this significant linear association disappears when we account for the effect of IQ on GPA. In other words, self-concept does not appear to be a useful linear predictor of GPA when considered as an additional predictor to IQ.