Problem Set #4-Key
Sonoma State University
Economics 317- Introduction to Econometrics

C1.1 For the regression equation wage = $\beta_0 + \beta_1$Education + U

(i) Describe the expected effects of education on wages (i.e., what is the expected sign of $\beta_1$).

(ii) Run the above regression. Are your results consistent with your expected effects in (i)?

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 526</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1179.73204</td>
<td>1</td>
<td>1179.73204</td>
<td>F( 1, 524) = 103.36</td>
</tr>
<tr>
<td>Residual</td>
<td>5980.68225</td>
<td>524</td>
<td>11.4135158</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>7160.41429</td>
<td>525</td>
<td>13.6388844</td>
<td>R-squared = 0.1648</td>
</tr>
</tbody>
</table>

| wage Coef. Std. Err. t P>|t| [95% Conf. Interval] |
|-----------------|------------|--------|---------------------|
| educ .5413593   | .053248    | 10.17  | 0.000 .4367534 .6459651 |
| _cons -.9048516| .6849678   | -1.32  | 0.187 -2.250472 .4407687 |

(iii) Show graphically the regression equation. Describe you results.

![Graph showing the regression equation and hourly wage vs. years of education.]

(iv) Use the $R^2$ and F-test to test for overall significance of the estimate regression. Explain each.

$R^2 = 0.1648$

F-test

$H_0: \beta_1 = \beta_2 = \ldots = \beta_n = 0$

$H_1: \beta_1 = \beta_2 = \ldots = \beta_n \neq 0$

Critical $F_{df1, df2, \alpha} = F_{1, 524, 0.05} = 3.8592662$

Sample $F = 103.36$

Reject $H_0$
(v) Use the three methods covered in class to test the coefficient on education for statistical significance. Be sure to formally state your hypothesis, use a 5% level of significance. Provide an explanation for each.

Use a one-sided hypothesis test. Why?

$H_0$: $\beta_1 \leq 0$
$H_1$: $\beta_1 > 0$

Using the t-test, the critical $t^* = 1.6477668$ while the sample $t = 10.17$, so reject $H_0$.

Using the p-value, $p = 0.000 < .05$, so reject $H_0$.

(vi) Construct a 95% confidence interval around the estimated coefficient $\hat{\beta}_1$. Explain.

$$\hat{\beta}_1^* = H_0 + t_{0.05,df} S_{\hat{\beta}_1}$$
$$1.6477668(.053248)$$
$$= .08774028$$

Sample $\beta_1 >$ Critical $\beta_1$ so reject $H_0$.

$\hat{\beta}_1 = .5413593 \pm 1.9645015(.053248)$
$.43675352 \text{ to } .64596508$
C1.2 For the regression equation Birth weight = β₀ + β₁Cigarettes + U

(i) Describe the expected effects of cigarettes on birth weight (i.e., what is the expected sign of β₁).

(ii) Run the above regression. Are your results consistent with your expected effects in (i)?

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs= 1388</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>13060.4194</td>
<td>1</td>
<td>13060.4194</td>
<td>F(1, 1386) = 32.24</td>
</tr>
<tr>
<td>Residual</td>
<td>561551.3</td>
<td>1386</td>
<td>405.159668</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>574611.72</td>
<td>1387</td>
<td>414.283864</td>
<td>R-squared = 0.0227</td>
</tr>
</tbody>
</table>

Adj R-squared = 0.0220

<table>
<thead>
<tr>
<th>bwght Coef.</th>
<th>Std. Err.</th>
<th>t</th>
<th>P&gt;t</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>cigs</td>
<td>-.5137721</td>
<td>.0904909</td>
<td>-5.68</td>
<td>0.000</td>
</tr>
<tr>
<td>_cons</td>
<td>119.7719</td>
<td>.5723407</td>
<td>209.27</td>
<td>0.000</td>
</tr>
</tbody>
</table>

(iii) Show graphically the regression equation. Describe your results.

(iv) Use the R² and F-test to test for overall significance of the estimate regression. Explain each.

R² = 0.0227

F-test

H₀: β₁ = β₂ = ... = βₙ = 0
H₁: β₁ ≠ β₂ ≠ ... ≠ βₙ ≠ 0

Critical F_{df1,df2,α} F_{1,1386,0.05} = 3.8481768

Sample F = 3.224

Reject H₀
(v) Use the three methods covered in class to test the coefficient on cigarettes for statistical significance. Be sure to formally state you hypothesis, use a 5% level of significance. Provide an explanation for each.

Use a one sided hypothesis test. Why?

H₀: β₁ ≥ 0
H₁: β₁ < 0

β₁* = H₀ - t_{α,df} S[β₁] 
-1.6459538(.0904909) 
= -.14894384

Sample β₁ < Critical β₁ so reject H₀

Using the t-test, the critical t* = -1.6459538 while the sample t = -5.68, so reject H₀.
Using the p-value, p = 0.000 < .05, so reject H₀.

(vi) Construct a 95% confidence interval around the estimated coefficient β₁. Explain.

Confidence Interval = \hat{β}_1 ± t_{α/2,df} S[β₁] 
-.5137721 ± 1.9616771(.0904909) 
-.69128602 to -.33625818
C1.3 For the regression equation Math Pass (Math4) = $\beta_0 + \beta_1\text{Expenditures Per Pupil} + U$

(i) Describe the expected effects of per pupil expenditures on the percentage of people who score satisfactorily on the mathematics test (i.e., what is the expected sign of $\beta_1$).

(ii) Run the above regression. Are your results consistent with your expected effects in (i)?

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 1823</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>705.779242</td>
<td>1</td>
<td>705.779242</td>
<td>F(1, 1821) = 1.77</td>
</tr>
<tr>
<td>Residual</td>
<td>724752.271</td>
<td>1821</td>
<td>397.996854</td>
<td>Prob &gt; F = 0.1831</td>
</tr>
<tr>
<td>Total</td>
<td>725458.05</td>
<td>1822</td>
<td>398.16578</td>
<td>R-squared = 0.0010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Adj R-squared = 0.0004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>725458.05</td>
<td>1822</td>
<td>398.16578</td>
<td>Root MSE = 19.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>math4</th>
<th>Coef.</th>
<th>Std. Err.</th>
<th>t</th>
<th>P&gt;t</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>expppp</td>
<td>-0.00057</td>
<td>0.000428</td>
<td>-1.33</td>
<td>0.183</td>
<td>-.0014095 .0002695</td>
</tr>
<tr>
<td>_cons</td>
<td>74.87011</td>
<td>2.272184</td>
<td>32.95</td>
<td>0.000</td>
<td>70.41375 79.32648</td>
</tr>
</tbody>
</table>

(iii) Show graphically the regression equation. Describe you results.

(iv) Use the $R^2$ and F-test to test for overall significance of the estimate regression. Explain each.

$R^2 = 0.0010$

F-test

$H_0: \beta_1 = \beta_2 = \ldots = \beta_n = 0$

$H_1: \beta_1 = \beta_2 = \ldots = \beta_n \neq 0$

Critical $F_{df1,df2,\alpha} F_{1,1821,.05} = 3.8465705$

Sample $F = 1.77$

Fail to Reject $H_0$
(v) Use the three methods covered in class to test the coefficient on per pupil expenditures for statistical significance. Be sure to formally state you hypothesis, use a 5% level of significance. Provide an explanation for each.

Use a one sided hypothesis test. Why?

\[ H_0: \beta_1 \leq 0 \]
\[ H_1: \beta_1 > 0 \]

\[ \beta_1^* = H_0 + t_{0.05, df_\beta_1} S_{\beta_1} \]
\[ = 1.6456908(0.000428) \]
\[ = .00070436 \]

Sample \( \beta_1 < \) Critical \( \beta_1 \) so fail to reject \( H_0 \)

Using the t-test, the critical \( t^* = 1.6456908 \) while the sample \( t = -1.33 \), so fail to reject \( H_0 \).

Using the p-value, \( p = 0.183 > 0.05 \), so fail to reject \( H_0 \).

(vi) Construct a 95% confidence interval around the estimated coefficient \( \beta_1 \). Explain.

Confidence Interval = \( \hat{\beta}_1 \pm t_{0.025, df_\beta_1} S_{\beta_1} \)
\[ = -.00057 \pm 1.9612676 (0.000428) \]
\[ = -.00140942 to .00026942 \]