5.1 (a) The 95% confidence interval for  is , that is .

(b) Calculate the *t*-statistic:

The *p*-value for the test  vs.  is

*p*-value

The *p*-value is less than 0.05, but larger than 0.01. So we can reject the null hypothesis at the 5% significance level, but not at the 1% significance level.

(c) The *t*-statistic is



The *p*-value for the test is

*p*-value

The *p*-value is larger than 0.10, so we cannot reject the null hypothesis at the 10%, 5% or 1% significance level. Because is not rejected at the 5% level, this value is contained in the 95% confidence interval.

(d) The 90% confidence interval for *β*0 is , that is, .

6.1. By equation (6.15) in the text, we know



Thus, the values of  for columns (1)–(3) are: 0.1648, 0.1817, 0.1843.

6.2. (a) Workers with college degrees earn $10.47/hour more, on average, than workers with only high school degrees.

(b) Men earn $4.69/hour more, on average, than women.

6.3. (a) On average, a worker earns $0.61/hour more for each year he ages.

(b) Sally’s earnings prediction is 0.11 + 10.44×1 − 4.56×1 + 0.61×29 = 23.68 dollars per hour.

Betsy’s earnings prediction is 0.11 + 10.44×1 − 4.56×1 + 0.61×34 = 26.73 dollars per hour. The difference is 3.05 $/hour (= 0.61× (34 − 29)).

6.4. (a) Workers in the Northeast earn $0.74 more per hour than workers in the West, on average, controlling for other variables in the regression. Workers in the Midwest earn $1.54 less per hour than workers in the West, on average, controlling for other variables in the regression. Workers in the South earn $0.44 less than workers in the West, controlling for other variables in the regression.

(b) The regressor *West* is omitted to avoid perfect multicollinearity. If *West* is included, then the intercept can be written as a perfect linear function of the four regional regressors.

(c) The expected difference in earnings between Juanita and Jennifer is

−0.44 − (−1.54) = $1.10/hour.

6.9. For omitted variable bias to occur, two conditions must be true: *X*1 (the included regressor) is correlated with the omitted variable, and the omitted variable is a determinant of the dependent variable. Since *X*1 and *X*2 are uncorrelated, the OLS estimator of *β*1 does not suffer from omitted variable bias.

**Empirical Exercise 5.3**

(a) Average birthweights, along with standard errors are shown in the table below. (Birthweight is measured in grams.)

|  |  |  |  |
| --- | --- | --- | --- |
|  | All Mothers | Non-smokers | Smokers |
|  | 3383 | 3432.1 | 3178.8 |
| SE() | 10.8 | 11.9 | 24.0 |
| *n* | 3000 | 2418 | 582 |

(b) The estimated difference is  = −253.2. The standard error of the difference is .

The 95% confidence for the difference is −253.2 ± 1.96×26.8 = (−305.9,−200.6).

(c) The estimated regression is

 = 3432.1 − 253.2*Smoker*

(11.9) (26.8)

(i) The intercept is the average birthweight for non-smokers (*Smoker* = 0). The slope is the difference between average birthweights for smokers (*Smoker* = 1) and non-smokers (*Smoker* = 0).

(ii) They are the same.

(iii) This the same as the confidence interval in (b).

(d) Yes − and we’ll investigate this more in future empirical exercises.