Economics 471: Econometrics

Department of Economics, Finance and Legal Studies
University of Alabama

Problem Set #2 – Answers

1. (a) The difference between β and $\widehat{\beta}$ is that β is the population parameter and is nonrandom (fixed), while $\widehat{\beta}$ is constructed from an estimator based on the random sample at hand and is random. There is no reason to believe that they are equivalent.

   (b) Both u and $\widehat{u}$ represent differences between the actual outcome and the casual relationship between the dependent and independent variables. However, u represents the difference between the actual outcome and the true casual relationship, while $\widehat{u}$ is the difference between the actual outcome and the estimated casual relationship. There is no reason to believe that they are equivalent as well.

   (c) The population model is the particular casual relationship that the researcher is interested in investigating, while the sample model is the estimate of the population model based off of the random sample collected by the researcher. There is no reason to believe that they are equivalent either.

2. The two assumptions are (i) $E(u) = 0$ and (ii) $E(ux) = 0$. These two assumptions are not necessary to make when deriving the OLS estimates, they merely are introduced to avoid using calculus to minimize the sum of squared errors (implicit assumptions). The first assumption is needed to identify the intercept. The second assumption is needed to identify the slope coefficient in our simple linear regression model and to guarantee that β represents the true casual effect of x on y.

3. β represents the true population parameter, while the remaining portion of the right hand side is the force that is inhibiting the researcher from fully learning about β. Here, since we do not know the errors, the second term is incalculable, which makes knowledge of β impossible. The goal of a researcher is to spicy a model and gather data in a way that makes the second term as close to zero as possible.

4. (a) The causality between cigarettes and income (assuming there is one) should be that income influences cigarette consumption. It can be argued that as income rises, health becomes more of a concern and so less smoking occurs. This scenario would lead to a negative slope coefficient.

   (b) $\#cigarettes = \alpha + \beta * income + u$

   (c) Average yearly income is \$19,304.83 and average number of cigarettes smoked a day is 8.69.

   (d) 498 out of 807 or roughly 62% do not smoke.

   (e) OLS estimations leads to

$$\widehat{cigarettes} = 7.14 + 0.00007986 * income$$

Here the slope coefficient is taken to mean that an increase of \$1,000 in yearly income results in 0.07986 more cigarettes smoked per day. This suggests that yearly income's economic impact on cigarette consumption is quite small.

(f) $R^2 = 0.002$ so roughly 1/5 of 1% of the total variation in cigarettes smoked per day is explained by income.

(g) $\epsilon_{income} = \frac{\partial cigatettes}{\partial income} * \frac{income}{cigarettes} \left(\approx \frac{\%\Delta cigatettes}{\%\Delta income} \right) = \widehat{\beta} * \frac{income}{cigarettes}$, evaluated at the averages we get $0.00007986 * \frac{19304.83}{8.69} = 0.1775$, which suggests that cigarettes are a normal good.

5. (a) As with the cigarette model determining the casual relationship is not clear cut. However, it would seem likely that total time spent working may influence total time spent sleeping rather than vice-versa.

(b) Average number of hours spent sleeping per work is 54.44 and the average number of hours spent working is 35.38. $\widehat{sleep} = 3586.38 - 0.1507 * work$; $R^2 = 0.1033$

(c) The intercept coefficient suggests that a person who does not work at all sleep on average 3586.38 or 59.77 hours per week. The slope estimates implies that, on average, an extra hour of work results in 9.04 fewer hours slept per week which is not even two hours less a night.