Economics 471: Econometrics

Department of Economics, Finance and Legal Studies
University of Alabama

Problem Set #6 – Answers

1. (a) $\widehat{sleep} = 3,840.83 - 0.163totwrk - 11.71educ - 8.70age + 0.128age^2 + 87.75male$

   (b) The coefficient on male is 87.75, so a man is estimated to sleep almost one and one-half hours more per week than a comparable woman. Further, $t_{male} = 87.75/34.33 \approx 2.56$, which is close to the 1% critical value against a two-sided alternative (about 2.58). Thus, the evidence for a gender differential is fairly strong.

   (c) The t-statistic on *totwrk* is $0.163/.018 \approx 9.06$, which is statistically significant. The coefficient implies that one more hour of work (60 minutes) is associated with $.163(60) \approx 9.8$ minutes less sleep.

   (d) We need to estimate the model without *age* and age_2. When *age* and age^2 are both in the model, *age* has no effect only if the parameters on both terms are zero.

   (e) The marginal effect is $\frac{\partial \widehat{sleep}}{\partial age} = -8.70 + 0.256age$. Note that the partial effect depends on the age of the individual.

2. (a) $\widehat{sat} = 1,028.10 + 19.30hsize - 2.19hsize^2 - 45.09female - 169.81black + 62.32female \cdot black$

   (b) The t statistic on $hsize^2$ is over four in absolute value, so there is very strong evidence that it belongs in the equation. The optimal point is obtained by finding the turnaround point; this is the value of *hsize* that maximizes (other things fixed): $19.3/(2 * 2.19) \approx 4.41$. Because *hsize* is measured in hundreds, the optimal size of graduating class is about 441.

   (c) This is given by the coefficient on female (since $black = 0$): nonblack females have SAT scores about 45 points lower than nonblack males. The t statistic is about –10.51, so the difference is statistically significant. (The large sample size certainly contributes to the statistical significance.)

   (d) Because $female = 0$, the coefficient on black implies that a black male has an estimated SAT score almost 170 points less than a comparable nonblack male. The t statistic is over 13 in absolute value, so we reject the hypothesis that there is no ceteris paribus difference.

   (e) We plug in $black = 1$, $female = 1$ for black females and $black = 0$ and $female = 1$ for nonblack females. The difference is therefore $-169.81 + 62.31 = 107.50$. Because the estimate depends on two coefficients, we cannot construct a t statistic from the information given. The easiest approach is to define dummy variables for three of the four race/gender categories and choose nonblack females as the base group. We can then obtain the t statistic we want as the coefficient on the black female dummy variable.

3. (a) The two signs that are pretty clear are $\beta_3 < 0$ (because *hsperc* is defined so that the smaller the number the better the student) and $\beta_4 > 0$. The effect of size of graduating class is not clear. It is also unclear whether males and females have systematically different GPAs. We may think that $\beta_6 < 0$, that is, athletes do worse than other students with comparable characteristics. But remember, we are controlling for ability to some degree with *hsperc* and *sat*.

   (b) $\widehat{colgpa} = 1.241 + 0.0569hsize + 0.00468hsize^2 + 0.0132hsperc + 0.00165sat + 0.155female + 0.169athlete$ Holding other factors fixed, an athlete is predicted to have a GPA about .169 points higher than a nonathlete. The t statistic $.169/.042 \approx 4.02$, which is significant.

   (c) With sat dropped from the model, the coefficient on athlete becomes about .0054 (se .0448), which is practically and statistically not different from zero. This happens because we do not control for SAT scores, and athletes score lower on average than nonathletes. Part (*b*) shows that, once we account for SAT differences, athletes do better than nonathletes. Even if we do not control for SAT score, there is no difference.

   (d) To facilitate testing the hypothesis that there is no difference between women athletes and women nonathletes, we should choose one of these as the base group. We choose female nonathletes. The estimated equation is

$$\begin{aligned}\widehat{colgpa} &= 1.396 + 0.0568hsize + 0.00467hsize^2 + 0.0132hsperc + 0.00165sat + 0.175female\_athlete \\ &\quad +0.013male\_athlete + 0.155male\_nonathlete\end{aligned}$$

   The coefficient on $fem\_athlete = female \times athlete$ shows that *colgpa* is predicted to be about .175 points higher for a female athlete than a female nonathlete, other variables in the equation fixed. The hypothesis that there is no difference between female athletes and female nonathletes is testing by using the t statistic on $female\_athlete$. In this case, t = 2.08, which is statistically significant at the 5% level against a two-sided alternative.

   (e) Whether we add the interaction female sat to the equation in part (*b*) or part (*d*), the outcome is practically the same. For example, when female sat is added to the equation in part (*b*), its coefficient is about .000051 and its t statistic is about .40. There is very little evidence that the effect of sat differs by gender.