Econ 312: Problem Set #3

Solutions

Question1. Sir Francis Galton, a cousin of James Darwin, examined the relationship between the height of children and their parents towards the end of the 19th century. It is from this study that the name "regression" originated. You decide to update his findings by collecting data from 110 college students, and estimate the following relationship:

$$\widehat{Studenth} = \underset{(7.2)}{19.6} + \underset{(0.10)}{0.73} \times Midparh,\ R^2 = 0.45,\ SER = 2.0$$

where *Studenth* is the height of students in inches, and *Midparh* is the average of the parental heights. Values in parentheses are heteroskedasticity robust standard errors. (Following Galton's methodology, both variables were adjusted so that the average female height was equal to the average male height.)

(a) Test for the statistical significance of the slope coefficient.

Answer: $H_0 : \beta_1 = 0$, t=7.30, the critical value for a two-sided alternative is 1.645. Hence we reject the null hypothesis

(b) If children, on average, were expected to be of the same height as their parents, then this would imply two hypotheses, one for the slope and one for the intercept.

(i) What should the null hypothesis be for the intercept? Calculate the relevant t-statistic and carry out the hypothesis test at the 1% level.
(ii) What should the null hypothesis be for the slope? Calculate the relevant t-statistic and carry out the hypothesis test at the 5% level.

Answer: $H_0 : \beta_0 = 0$, t=2.72, for $H_1 : \beta_0 \neq 0$, the critical value for a two-sided alternative is 2.58. Hence we reject the null hypothesis in (i). For the slope we have $H_0 : \beta_1 = 1$, t=-2.70, for $H_1 : \beta_1 \neq 1$, the critical value for a two-sided alternative is 1.96. Hence we reject the null hypothesis in (ii).

(c) Can you reject the null hypothesis that the regression R^2 is zero?

Answer: For the simple linear regression model, $H_0 : \beta_1 = 0$ implies that $R^2 = 0$. Hence it is the same test as in (a).

(d) Construct a 95% confidence interval for a one inch increase in the average of parental height.

Answer: $(0.73 - 1.96 \times 0.10,\ 0.73 + 1.96 \times 0.10) = (0.53,\ 0.93)$.

Question.2 A researcher plans to study the causal effect of police on crime using data from a random sample of U.S counties. He plans to regress the county's crime rate on the (per capita) size of the county's police forces.

(a) Explain why this regression is likely to suffer from omitted variable bias. Which variables would you add to the regression to control for important variables?

Answer: There are other important determinants of a country's crime rate, including demographic characteristics of the population.

(b) Use you answer to (a) and the expression for omitted variable bias to determine whether the regression will likely over-or underestimate the effect of police on the crime rate.

Answer: Suppose that the crime rate is positively affected by the fraction of young males in the population, and that counties with high crime rates tend to hire more police. In this case, the size of the police force is likely to be positively correlated with the fraction of young males in the population leading to a positive value for the omitted variable bias so that $\hat{\beta}_1 > \beta_1$.