

Empirical Exercises for Chapter 4

In this exercise you will investigate the relationship between the number of completed years of education for young adults and the distance from their high school to the nearest four-year college. Proximity to college lowers the cost of education, so that students who live closer to a four-year college should, on average, complete more years of schooling. But how much will completed years of education increase if a college is built within 10 miles of a student's high school instead of 20 miles?

You will examine this question using data on a random sample of high school seniors, contained in the data file **CollegeDistance** (in Excel and Stata formats). These same students were interviewed as high school seniors in 1980, then reinterviewed in 1986 to determine how many years of education they had completed. A detailed description is given in **CollegeDistance_Description**, available on the Web site.

Use these data to answer the following questions.

1. Run a regression of years of completed education (ED) on distance to the nearest college ($Dist$), where $Dist$ is measured in 10's of miles. (For example, $Dist = 2$ means that the distance is 20 miles.) What is the estimated intercept? What is the estimated slope? Use the estimated regression to answer the question "How does the average value of years of completed schooling change when colleges are built close to where students go to high school?"
2. Is the estimated regression slope coefficient statistically significant? That is, can you reject the null hypothesis $H_0: \beta_1 = 0$ versus a two-sided alternative at the 10%, 5%, or 1% significance level? What is the p -value of the test?
3. Construct a 95% confidence interval for the slope coefficient.
4. Bob's high school was 20 miles from the nearest college. Predict Bob's years of completed education using the estimated regression.
5. Does distance to college explain a large fraction of the variance in educational attainment across individuals? Explain.
6. What is the value of the standard error of the regression? What are the units for the standard error (meters, grams, years, dollars, cents, bananas, etc.)?
7. Repeat question (1) using only the data on females.
8. Repeat question (1) using only the data on males.
9. Is the effect of distance on completed years of education different for men than for women? (Hint: Assume that the sample of men and women were independently drawn.)