

EC1630 Spring 2012

Homework 1

Due February 15 (beginning of class)

1. A national telephone survey found the following: 248 of the 460 adults questioned supported a government-administered insurance plan, while the others said they were opposed. Let p denote the fraction of all adults who supported the plan at the time of the survey, and let \hat{p} be the fraction of survey respondents who supported the plan.
 - (a) If survey respondents are i.i.d draws from the population of adults, show that \hat{p} is an unbiased estimator of p .
 - (b) How would you calculate the standard error of the estimator \hat{p} ? (Hint: show that $\text{Var}[\hat{p}] = p(1 - p)/n$, where n is the number of draws.)
 - (c) What is the p-value for the test $H_0 : p = 0.5$ vs. $H_1 : p \neq 0.5$? Will you reject the null at the 5% significance level?
 - (d) What is the p-value for the test $H_0 : p = 0.5$ vs. $H_1 : p > 0.5$? Will you reject the null at the 5% significance level?
 - (e) Construct a 95% confidence interval for p (two-sided).
2. Using the data set `CollegeDistance.dta` (see Empirical Exercise 4.3 and 5.3 (**SW**)), run a regression of years completed education (ed) on distance to the nearest college ($dist$) and carry out the following exercises.
 - (a) 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 associated with the coefficient's t-statistic?
 - (b) Construct a 95% confidence interval for the slope coefficient.
 - (c) Run the regression using data only on females and repeat (b) (to obtain $\hat{\beta}_{w,1}$).
 - (d) Run the regression using data only on males and repeat (b) (to obtain $\hat{\beta}_{m,1}$).
 - (e) Is the effect of distance on completed years of education different for men than for women?
(Hint: The standard error of $\hat{\beta}_{m,1} - \hat{\beta}_{w,1}$ is given by $\sqrt{[SE(\hat{\beta}_{m,1})]^2 + [SE(\hat{\beta}_{w,1})]^2}$.)
 - (f) Make a scatter plot of ed against $dist$ and display the fitted regression line. Does $dist$ explain a lot of the variation in ed ? Explain. (Hint: Also look at the regression output obtained in (a).)