Professor Joyce Jacobsen

Economics 385
Spring Semester 2013-14
March 5, 2014

Midterm

Each problem is weighted equally. In order to get full credit, you must show any calculations used to arrive at your answers and answer the questions completely.

1. Which of the following are fixed/nonstochastic and which are random/stochastic?

$$Y,\ X,\ u,\ \hat{u},\ \beta,\ \hat{\beta},\ \sigma^2,\ \hat{\sigma}^2,\ \mathrm{Var}(\hat{\beta}),\ \hat{\mathrm{Var}}(\hat{\beta}),\ R^2,\ \bar{R}^2$$

2. You have data on two variables, Y and X, and plan to estimate $Y = \beta_0 + \beta_1 X$. In each of the following cases, if you are given the following information but nothing else, explain which of the two samples you prefer to use for your estimation (or if you are indifferent), and why:

   a) Sample 1 has twice as many observations as Sample 2.

   b) Samples 1 and 2 have the same number of observations, but S_{XX} is twice as big in Sample 2 as in Sample 1.

   c) Samples 1 and 2 have the same number of observations, but S_{YY} is twice as big in Sample 2 as in Sample 1.

   d) Sample 1 has twice as many observations as Sample 2, S_{XX} is twice as big in Sample 1 as in Sample 2, and S_{YY} is twice as big in Sample 1 as in Sample 2.

3. For applying OLS to the classical linear model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$:

   a) If the variance of X_1 increases, does the variance of $\hat{\beta}_0$ increase, decrease, or is it unaffected?

   b) If $\hat{\beta}_0$ is negative, does $\bar{Y}$ have to be negative as well? Explain why or why not.

   c) If the covariance of $\hat{\beta}_1$ and $\hat{\beta}_2$ is positive and both X_1 and X_2 are on average positive, is the covariance of $\hat{\beta}_0$ and $\hat{\beta}_1$ positive, negative, or of undeterminable sign from this information?

4. You buy a mine (the rights to a hole in the ground). The prospector who sold you the mine told you that at least ten percent of the ore is gold. Ore comes in the form of uniformly-sized pebbles, each of which is either all gold or all not gold.

   a) You randomly pull out a pebble and it is not gold. What is your best estimate of what percent of the ore is gold, conditional on believing that the seller told you the truth?

   b) You randomly pull out ten pebbles and none of them are gold. Now how much less sure are you that the seller told you the truth (as opposed to the alternative that the mine is worthless)? If you randomly pull out a thousand pebbles and none of them are gold, how much less sure are you?

   c) Should you ever simply believe someone's statement, or should you base your assessment of the truth or falsity of a statement on the relevant empirical evidence? How does econometrics help us reconcile these two positions?

5. Suppose the classical linear model applies to $y_i = \beta_0 + \beta_1 x_i + u_i$. A friend of yours suggests estimating β_1 as:

$$\frac{1}{n-2}\sum_{i=3}^{n}\frac{y_i - y_2}{x_i - x_2}$$

a) Show that this is an unbiased estimator of β_1

b) Should you follow your friend's suggestion? Why or why not?

Your friend then suggests evaluating whether this estimator is good or not by using the leftover observation (x_1, y_1) first to calculate $\hat{y}_1$, then to calculate the squared prediction error for this leftover observation, and then to use this as an estimate of the mean squared error (MSE) for this estimation method.

c) Would this be a good estimate of the MSE? Why or why not?

d) Would you expect this calculation of MSE to be greater or smaller than the OLS MSE?

6. Suppose the classical linear model applies to $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + u$. Suppose a (different) friend of yours tells you that $\beta_1 = 2\beta_2 - \beta_3$

a) Explain what regression to run to find the OLS estimates that incorporate this information from your friend.

b) How would you test formally whether or not your friend is telling the truth?

7. If you were to run a regression equation where earnings = f(education, age), give an example of:

a) an omitted variable that is likely biasing the estimate of the coefficient on age upward

b) an omitted variable that is likely biasing the estimate of the coefficient on education downward.

c) an omitted variable that is appropriately omitted from this equation (i.e., is irrelevant)

8. For each of the following statements, state whether it is true, false, or uncertain, and explain your answer.

a) $(X'X)^{-1}$ is idempotent.

b) R^2 can never be greater than 1.

c) $\bar{R}^2$ can only be appropriately used to choose between nested models.