Professor Joyce Jacobsen

Economics 385
Spring Semester 2010-11
March 4, 2011

Midterm

Each problem is weighted equally out of 100 points. In order to get full credit, you must show any calculations used to arrive at your answers and completely answer the questions.

1. Explain each of the following concepts as it has been developed in the course:

   β, $\hat{\beta}$, σ^2, $\hat{\sigma}^2$, Var($\hat{\beta}$), estimatedVar($\hat{\beta}$)

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

   a) Explain why $\hat{\beta}_0$ cannot be derived independently from $\hat{\beta}_1$ and $\hat{\beta}_2$.

   b) Why is the fitted line forced through the means of the data—that is, contains the point ($\bar{X}_1, \bar{X}_2, \bar{Y}$)?

3. Explain why it is generally considered worse to omit a true variable from the fitted model than it is to add in an extraneous variable.

4. You observe three numbers, 3, 6, and 8, drawn independently from an unknown distribution.

   a) If this distribution were uniform, what is your best guess as to the parameter of the distribution (namely the range)?

   b) If this distribution were normal, what is your best guess as to the parameters of the distribution (namely the mean and variance, or mean and standard deviation if you prefer)?

   c) Since you don't actually know what type of distribution this is, you and your friend argue. You pick one and ze picks the other. Explain compellingly to your friend why you picked yours.

   d) Suppose you and your friend see another twenty numbers drawn independently from this unknown distribution. Explain how that additional information would make it clearer which of you was right. Explain also how, in either case (uniform or normal), how that additional information would make it clearer what the true parameter or parameters of the distribution are.

5. Suppose you have the following data from 100 observations on the classical linear model
   $Y = \alpha + \beta X + u$, $V(u) = \sigma^2$:

   $\Sigma X = 200$ $\Sigma X^2 = 500$ $\Sigma XY = 400$ (note these are all levels, not deviations)
   $\Sigma Y = 100$ $\Sigma Y^2 = 10{,}300$

   A number *z* is calculated as $2\alpha + 9\beta$. A number *q* is formed by throwing 10 fair dice and adding the total number of spots appearing. A contest is being held to guess $W = z + q$ with contestants being rewarded (or penalized) according to the formula $p = 60 - (W - W^*)^2$ dollars, where W^* is the contestant's guess.

   a) What would your guess be if you wanted to make the expected value of your guess equal to the expected value of *W*?

   b) What is the expected payoff of this guess? (Hint: get an expression in terms of σ^2)

6. Suppose the classical linear model applies to $Y = \alpha + \beta X + u$. Your friend (yes, the same annoying one as in question 4) multiplies all your X values by 3. If the old Y are regressed on the new X, what can you say about the expected values of your estimates of α and β?

7. Consider applying OLS to a consumption function $C = \alpha + \beta Y$ and to the corresponding saving function $S = \gamma + \delta Y$ where for all observations $Y = C + S$.

   a) Show that $\hat{\delta} = 1 - \hat{\beta}$

   b) The sum of squared residuals is the same for each regression. True, false, or uncertain? Explain.

   c) The $\bar{R}^2$s are the same for each regression. True, false, or uncertain? Explain.

8. Imposing a linear constraint on a regression will raise R^2 if the constraint is true and lower R^2 if it is false. True, false, or uncertain? Explain.

9. Your friend (this time a different one) asserts that instead of using OLS to estimate the parameters of the simple linear model, that researchers should find the $\hat{\beta}_0$ and $\hat{\beta}_1$ that minimize $\sum |Y - \beta_0 - \beta_1 X|$

   a) Argue with him as to why OLS is preferable to his method.

   b) Now switch sides and agree with him. On what grounds do you agree with him?

   c) We often use the criterion of low mean squared error (MSE) to decide when one estimator is preferable to another. Explain why that criterion tends to favor choosing OLS over other estimators. Can you come up with an alternative criterion that would tend to favor choosing your friend's estimator over other estimators?

10. A study on unemployment in the British interwar period produced the following regression equation:

$$U = 5.19 + 18.3(B/W) - 90.0(\log Q - \log Q^*) \; ; \; R^2 = 0.8, \text{ SE of the regression} = 1.9$$

sample period 1920-1938 (n = 19)
where U = unemployment rate,
B/W = ratio of unemployment benefits to average wage,
Q = actual output,
Q* = trend predicted output,
so logQ – logQ* captures unexpected changes in aggregate demand

The authors concluded that the high benefit levels were partly responsible for the high rates of unemployment. Critics of this study argued that when the single observation for 1920 is dropped, the results change dramatically. The equation then is:

$$U = 7.9 + 12.9(B/W) - 87.0(\log Q - \log Q^*) \; ; \; R^2 = 0.82, \text{ SE of the regression} = 1.7$$

sample period 1921-1938 (n = 18)

Test whether the results are significantly different from each other.