14.381, Fall 2006
Problem Set 6
Due: Dec. 5, 2006 (in class)

1. **Linearity and Unbiasedness.** This problem asks you to think about the meaning of linearity and unbiasedness while offering practice with “thrilling" algebraic manipulation.

   Consider the regression model:

   $$y_t = x_t'\beta + \varepsilon_t,\ \ (t = 1, 2, \ldots, T) \tag{1}$$

   where $E[\varepsilon|x] = 0$ and $E[\varepsilon\varepsilon'|x] = \sigma^2 I$. Let $\bar{x} \equiv \frac{1}{T}\sum_{t=1}^{T} x_t$ and define $\bar{y}$ similarly.

   (a) Consider the estimator $\hat{\beta} \equiv \frac{\bar{y}}{\bar{x}}$. Show that $\hat{\beta}$ is linear and unbiased. Calculate its variance and compare it to that of the OLS estimator. Suppose you're working with microdata and cannot get individual data but only group averages. What does the preceding result tell you? (For an application of this idea, see Josh Angrist, “Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants,” Econometrica, 1998.)

   (b) Suppose that you decide to use only the first $\tau < T$ observations from your sample and do OLS. Show that this estimator $\tilde{\beta}$ is linear and unbiased but not of minimum variance.

   (c) Suggest a minimum variance, but not necessarily unbiased, estimator (Hint: this may feel like a bit of a trick question after you find the answer).

2. **Violating the sphericality assumption**. In this question you'll work through what happens when you don't have sphericality.

   Consider the regression model:

   $$y_t = x_t'\beta + \varepsilon_t,\ \ (t = 1, 2, \ldots, T) \tag{2}$$

   where $E[\varepsilon|x] = 0$ and $E[\varepsilon\varepsilon'|x] = \Sigma$, with Σ a positive definite matrix. Let $\hat{\beta}$ be the OLS estimator of β.

   (a) Is $\hat{\beta}$ unbiased? Compute its covariance matrix.

   (b) Let a be a fixed $K \times 1$ vector. Is $a'\hat{\beta}$ and unbiased estimator of $a\beta$? Compute $var\,(a'\beta)$.

3. **Review of Materials from 14.381 Part I + Some Key derivations one needs to know.**

   (a) Prove: if $Z \sim N(\mu, \Sigma)$ and A is a fixed, conformable matrix, then

   $$AZ \sim N(A\mu, A\Sigma A')$$

   Note: there are three parts to this question – you have to show that AZ has mean $A\mu$, variance $A\Sigma A'$ and that AZ has a normal distribution. Showing the normality part is very difficult; it is fine to show for an invertible A or for scalar a and z.

(b) Prove: if $Z \sim N(0, I)$ and A is a symmetric, idempotent matrix, then

$$Z'A'Z \sim \chi^2_J$$

where

$$J = rank\,(A)$$

(c) Let P_X be the X-projection matrix as defined in class, and $M_X = I - P_X$ be the corresponding annihilator matrix. Prove that

$$rank\,(P_X) = K$$

and

$$rank\,(M_X) = N - K$$

(d) Independence and uncorrelatedness.

1. Prove that if two random vectors are independent, they are uncorrelated.
2. Give a counterexample to the converse. That is, give an example of two uncorrelated RVs that are not independent. You can use random variables (unidimensional) if you want.
3. Prove that the converse does hold for jointly normal random vectors That is, if X and Y are jointly normally distributed and uncorrelated, then X and Y are independent. (Hint: review the definition of independence and stare hard at the formula for the joint density of jointly normal random vectors.)

(e) Let

$$X_n \sim t(n)$$

and

$$X \sim N(0, 1)$$

Prove that as $n \to \infty$

$$X_n \underset{d}{\to} X$$

Hints: write down the definition of a $t(n)$. What is the relationship between a χ^2_n random variable and a χ^2_1 random variable? Look up the mean of a χ^2_1 random variable. Ponder the deep meaning of the Law of Large Numbers.

(f) Prove that under GM1-5:

1. $\hat{\beta}|X \sim N(\beta, V)$, where $V = \sigma^2 (X'X)^{-1}$
   We were able to prove that $E\left[\hat{\beta}|X\right] = \beta$ and $V\left[\hat{\beta}|X\right] = \sigma^2 (X'X)^{-1}$ without imposing the assumption that

$$GM5 : \varepsilon|X \sim N\left(0, \sigma^2 I\right)$$

   So why should we bother?
2. $\left(\hat{\beta}_j - \beta\right) / \sqrt{V_{jj}} \sim N(0, 1)$
   Why is this result not incredibly useful on its own?

3. $\dfrac{(N-K)\,S^2}{\sigma^2} \sim \chi^2_{N-K}$

4. $\hat{\beta}$ and S^2 are independent
   Why is this result so suprising? (Borat type questions.)
   Why is it so obvious once you get it? (Hint: same as the answer to the previous question.)
   Why is it so important?
   How robust is it to our assumptions?

5. $\left(\hat{\beta}_j - \beta\right)/\sqrt{\hat{V}_{jj}} \sim t_{N-K}$, where $\hat{V}_{jj}$ is the jjth entry in $\hat{V} = S^2\,(X'X)^{-1}$ (Note that it is a little dangerous to write

$$\hat{V}_{jj} = S^2\,(X'X)^{-1}_{jj}$$

It is true that

$$\hat{V}_{jj} = S^2\left((X'X)^{-1}\right)_{jj}$$

but it is not true that

$$\hat{V}_{jj} = S^2\left((X'X)_{jj}\right)^{-1}$$

(can you see the difference) and it is not always clear which of these you mean by

$$\hat{V}_{jj} = S^2\,(X'X)^{-1}_{jj}$$

However, because this is confusing, it is standard notation.
Wait a minute: we had basically the same thing in part (2), and we didn't have to worry about degrees of freedom or kegs of Guinness or anything like that. What's so great about this result? Why bother with all the math in parts (3) and (4)?

4. **Reverse Regression.**

Consider the two models:

$$y_t = \alpha + x_t\beta + u_t \tag{3}$$
$$x_t = a + y_t b + v_t \tag{4}$$

for $t = 1, 2, \ldots, T$, where $E[u|\,x] = 0$, $E[uu'|\,x] = \sigma_u^2 I$, $E[v|\,y] = 0$, and $E[vv'|\,y] = \sigma_v^2 I$. Let $\hat{\beta}$ be the OLS estimator for β in model (3) and similarly $\hat{b}$ for b in (4).

(a) Suppose you find that $\hat{\beta} > 0$. Compare $\hat{\beta}$ and $\hat{b}$. Note: this problem may be easier to do if you drop the α and a, but you have to justify doing this.

(b) Compare the R^2 you get from estimating each model.

5. **Simulation and Hypotheses Testing**. *You'll conduct a hypotheses test on a simulated regression model and compare your empirical rejection percentage to the intended size of your test. Here, in a canonical model, you should find that they match very closely; however, when we start looking at violations of the GM assumptions, things*

will often go haywire. Simulations offer a great way to see if you're testing what you think you're testing. Bertrand, Duflo, and Mullinathan's "How Much Should We Trust Differences-in-Differences Estimates?" (QJE, 2004) offers a nice example of this.

Starting with the code[1] from question 6 of Problem Set 5, for each iteration and value of I, test the hypothesis $\beta = 2$ at a size of $\alpha = 0.05$. Record the number of times you reject. How does your rejection percentage compare to α? Should this vary with I? Hints:

(a) You can record rejections in either a counter or in a vector of 1's and 0's which you can then sum to get totals.

(b) Remember that the critical value for Student's t-distribution when the sample size, I, is 10 is not 1.96. You may want to use the "$tinv\left(1 - \frac{a}{2}, n\right)$" command in Matlab that returns the critical values of a t-distribution, where a denotes the size of the two-sided test and n the sample size.

(c) The sample code in the Matlab introduction may be helpful.

(d) You should only have to add about six lines of code to what you did in the previous problem set.

6. **Simple Regression Testing and Welcome to STATA.** *This problem asks you estimate the parameters of a production function and test some hypotheses about it. The point here is to get you to think about the meaning of a regression and the tests you are conducting. You can use STATA, or the like for the programming, but we'll be able to offer you more constructive feedback if you stick with STATA. For those of you completely unfamiliar with STATA, Chris Smith's tutorial is a good place to start. You can find a link to it at* http://web.mit.edu/econ-gea/computer.htm. You can also refer to a handout provided by Raymond Guiteras for more STATA commands of regressions. *Please submit a printout of your results, your do-file,* ***and*** *an explanation of what you did (embedded in the do-file would be great).*

You have a dataset on the output, labor, and capital of 27 firms in the primary metals industry (SIC 33). Each observation represents one firm. You can download the data from the course website. The data is available both in a text file (dataPset6Qn6.txt) and in a STATA data file (dataPset6Qn6.dta). Use either one at your convenience.

(a) Estimate a Cobb-Douglas production function of the form

$$\ln Q_i = \beta_1 + \beta_2 \ln L_i + \beta_3 \ln K_i + \varepsilon_i.$$

(b) Test the hypothesis of constant returns to scale. At what level could you reject? STATA kindly names the command you'll want to use *test*.

(c) The Translog model

$$\ln Q_i = \beta_1 + \beta_2 \ln L_i + \beta_3 \ln K_i + \beta_4 \frac{\ln^2 L_i}{2} + \beta_5 \frac{\ln^2 K_i}{2} + +\beta_6 \ln L_i \ln K_i + \varepsilon$$

[1] You can use Matlab or any other programming language of your choice.

offers greater flexibility than the Cobb-Douglas and has been often used for empirical estimation of production function. Estimate a Translog production function for the data you have.

(d) What restrictions on the Translog model give back the Cobb-Douglas specification? Test this restriction. At what level could you reject?

(e) In the Translog model, what is the mean value in the sample of the elasticity of output with respect to capital, $\frac{\partial \ln Q_i}{\partial \ln K_i}$, evaluated at the parameter estimates you calculated in (c)?

(f) Calculate the correlation between $\ln K_i$ and $\ln Q_i$. In the Cobb-Douglas model you calculated in (a), what is the effect of capital on output? Explain in simple terms why these two answers aren't the same.

(g) In the Cobb-Douglas model from (a), test the income distribution hypothesis that β_2 equals the share of labor in gross revenue, 0.80. Can you reject at the 5% level?

(h) Bonus: Power Curves. If the true β_2 were 0.80, how often would you reject in the above test? What if it were 0.60? Graph the probability of rejecting $H_0 : \beta_2 = 0.80$ as a function of β_2 for a test of size 5%.