14.381, Fall 2006
Problem Set 7
Due: Dec. 12, 2006 (in class)

1. **Assessing Consistency.** *This problem highlights the differences between unbiasedness and consistency.*

   Suppose you want to estimate the parameters of the model

   $$y_i = \beta_1 + \frac{1}{\beta_2} x_i + \varepsilon_i$$

   under the standard assumptions for OLS. You run OLS on the regression

   $$y_i = \delta_1 + \delta_2 x_i + \varepsilon_i$$

   to obtain $\hat{\delta}_2$ and then use the estimator $\hat{\beta}_2 = \frac{1}{\hat{\delta}_2}$.

   (a) Is $\hat{\beta}_2$ unbiased?

   (b) Is $\hat{\beta}_2$ consistent? If so, give a proof and state precisely the assumptions needed.

   (c) Derive the asymptotic distribution of $\hat{\beta}_2$.

   (d) Under the additional assumption that ε is normal, what is the maximum likelihood estimator of β_2?

2. **Assessing Consistency II.** *This problem asks you to assess consistency in the context of a dataset with missing observations but available proxies for the missing data. It's a nice example of how to check consistency and also offers some insight on how to think about data difficulties.*

   You are estimating the relationship

   $$y_i = \alpha + \beta x_i + \varepsilon_i$$

   where the standard OLS assumptions are again satisfied. Unfortunately, data on x are missing for some observations; however, data on a proxy z for x are always available. The proxy satisfies $x = z + \eta$, where η is uncorrelated with both z and ε.

   (a) Is OLS omitting observations where x is missing consistent?

   (b) Is OLS using x when it is available and z otherwise consistent?

   (c) Would the usual OLS standard error formula be correct for the estimator in (b)? What if $\beta = 0$?

   (d) Which of the above estimation proceedures would you prefer for testing the null hypothesis $\beta = 0$ and why?

3. **Heteroscedasticty-Robust Standard Errors and Testing Procedures.** *This problem requires you to repeat part (a) – (b) of Question 6 in Problem Set 6 under the assumption of heteroscedasticity. How does heteroscedasticity affect the following estimates and hypothesis tests?*

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

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

Report robust standard errors. Compare the results to the ones obtained using non-robust standard errors.

(b) Test the hypothesis of constant returns to scale, using robust estimates of variance matrices. At what level could you reject? Compare your results to the results obtained using non-robust version.

4. **Small Sample Inference Beyond Normality.** *Read Victor's "Finite Sample Inference Beyond Normality" handout (finitesample.pdf) on the Stellar site. Do the empirical exercise from Example 3. It's based on Temin (2005) which is also available on the course website.*

5. **A Quick Review of the Course:** Consider the model, where $Y = X\beta + \epsilon$, where for each t, $\epsilon_t \sim \sigma(e_t - 1)$, where e_t is standard exponential variable such that $E[e_t] = 1$ and $Var[e_t] = 1$. Assume that X are independent of ϵ. Suppose that (x_t, ϵ_t) are i.i.d. across t.

(a) Do Gauss-Markov assumptions hold for this model?

(b) Consider the least squares estimator $\hat{\beta}$. Compute $E[\hat{\beta}|X]$ and $Var[\hat{\beta}|X]$. Is $\hat{\beta}$ normally distributed in finite samples, conditional on X?

(c) Carefully, but briefly, explain the label "BLUE". Is OLS BLUE in this set-up?

(d) Consider estimating the following effect

$$E[y_t|x_t = x''] - E[y_t|x_t = x'] = (x'' - x')'\beta$$

Give an economic example where such an effect might be of interest. Is $(x'' - x')'\hat{\beta}$ BLUE (=Minimum Variance Linear Unbiased Estimator) for this effect? Why or why not?

(e) (Bonus). Is OLS the BUE (best unbiased estimator) in this model? A brief answer suffices.

(f) What is the large sample distribution of $\hat{\beta}$? Make any additional primitive assumptions you might need. [Note: high level assumptions will receive partial credit.]

(g) Construct a consistent estimator for the large sample variance of $\hat{\beta}$. Prove its consistency by making any additional assumptions you need.

(h) Suppose we want to test the null hypothesis $H_0 : \beta_j = 0$ vs $H_A : \beta_j < 0$. Construct a t-statistic for testing this hypothesis. Derive its limit distribution and describe how to select critical value for this test to maintain the level of significance equal to 5%.

(i) Suppose the sample size $n = 6$. Do you expect the large sample distribution to be a good approximation to the exact distribution of the t-statistic in question h.. Discuss how to get the exact distribution of the t-statistic. How would you generate p-values (or critical values) for checking the hypothesis of question h. that would be valid even for $n = 6$?