Econometrics 710
Final Exam, Spring 2008

1. Take the model

$$\begin{aligned} y_i &= z_i'\beta + e_i \\ E(x_i e_i) &= 0 \\ E\left(e_i^2\right) &= \sigma^2 \end{aligned}$$

Describe an estimator for σ^2.

2. You are reading a paper, and it reports the results from two nested OLS regressions:

$$y_i = x_{1i}'\tilde{\beta}_1 + \tilde{e}_i \qquad (1)$$
$$y_i = x_{1i}'\hat{\beta}_1 + x_{2i}'\hat{\beta}_2 + \hat{e}_i \qquad (2)$$

Some summary statistics are reported:

Regression (1)	Regression (2)
$R^2 = .20$	$R^2 = .26$
$\sum_{i=1}^n \tilde{e}_i^2 = 106$	$\sum_{i=1}^n \hat{e}_i^2 = 100$
# of coefficients=5	# of coefficients=8
$n = 50$	$n = 50$

You are curious if the estimate $\hat{\beta}_2$ is statistically different from the zero vector. Is there a way to determine an answer from this information? Do you have to make any assumptions (beyond the standard regularity conditions) to justify your answer?

3. Your model is

$$\begin{aligned} y_i^* &= x_i'\beta + e_i \\ E(e_i \mid x_i) &= 0 \end{aligned}$$

However, y_i^* is not observed. Instead only a capped version is reported. That is, the dataset contains the variable

$$y_i = \begin{cases} y_i^* & \text{if} \quad y_i^* \leq \tau \\ \tau & \text{if} \quad y_i^* > \tau \end{cases}$$

Suppose you regress y_i on x_i using OLS. Is OLS consistent for β? Describe the nature of the effect of the mis-measured observation on the OLS estimate.

4. The data $\{y_i, x_i, w_i\}$ is from a random sample, $i = 1, ..., n$. The parameter β is estimated by minimizing the criterion function

$$S(\beta) = \sum_{i=1}^{n} w_i \left(y_i - x_i'\beta\right)^2$$

That is $\hat{\beta} = \text{argmin}_\beta \, S(\beta)$

(a) Find an explicit expression for $\hat{\beta}$

(b) What population parameter is $\hat{\beta}$ estimating? (Be explicit about any assumptions you need to impose. But don't make more assumptions than necessary.)

(c) Find the probability limit for $\hat{\beta}$ as $n \to \infty$

(d) Find the asymptotic distribution of $\sqrt{n}\left(\hat{\beta} - \beta\right)$ as $n \to \infty$.

5. Suppose a PhD student has a sample $(y_i, x_i, z_i : i = 1, ..., n)$ and estimates by OLS the equation

$$y_i = z_i\hat{\alpha} + x_i'\hat{\beta} + \hat{e}_i$$

where α is the coefficient of interest and he is interested in testing $H_0 : \alpha = 0$ against $H_1 : \alpha \neq 0$. He obtains $\hat{\alpha} = 2.0$ with standard error $s(\hat{\alpha}) = 1.0$ so the value of the t-ratio for H_0 is $t_n = \hat{\alpha}/s(\hat{\alpha}) = 2.0$. To assess significance, the student decides to use the bootstrap. He uses the following algorithm:

(a) Samples (y_i^*, x_i^*, z_i^*) randomly from the observations. (Random sampling with replacement). Creates a random sample with n observations.

(b) On this pseudo-sample, estimates the equation

$$y_i^* = z_i^*\hat{\alpha}^* + x_i^{*\prime}\hat{\beta}^* + \hat{e}_i^*$$

by OLS and computes standard errors, including $s(\hat{\alpha}^*)$. The t-ratio for H_0, $t_n^* = \hat{\alpha}^*/s(\hat{\alpha}^*)$ is computed and stored.

(c) This is repeated $B = 9999$ times.

(d) The 95% empirical quantile $\hat{q}_{.95}^*$ of the bootstrap absolute t-ratios $|t_n^*|$ is computed. It is $\hat{q}_{.95}^* = 3.5$.

(e) The student notes that while $|t_n| = 2 > 1.96$ (and thus an asymptotic 5% size test rejects H_0), $|t_n| = 2 < \hat{q}_{.95}^* = 3.5$ and thus the bootstrap test does not reject H_0. As the bootstrap is more reliable, the student concludes that H_0 cannot be rejected in favor of H_1.

Question: Do you agree with the student's method and reasoning? Do you see an error in his method?