PS #6 (Due 27 Feb, 2019)

1. Write an R code and a Python code to implement the following:
   (a) generate 5 i.i.d. observations from the uniform random variable on [-0.5,0.5] and compute the sample average.
   (b) Repeat (a) 1000 times, using the new observations each time.
   (c) Draw the histogram of the 1000 averages.
   (d) Do (a)–(c) using 20 observations from the same random variables.
   (e) Do (a)–(c) using 80 observations from the same random variables.
   (f) If you draw the three histograms on the same scale, what do you see? Explain what you observe using the LLN.

2. Write an R code and a Python code to implement the following:
   (a) generate 5 i.i.d. observations from the uniform random variable on [-0.5,0.5] and compute the sample average.
   (b) Repeat (a) 1000 times, using the new observations each time.
   (c) Draw the histogram of the 1000 averages time square root of 5.
   (d) Do (a)–(c) using 20 observations from the same random variables, except this time multiply by square root of 20.
   (e) Do (a)–(c) using 80 observations from the same random variables, except this time multiply by square root of 80.
   (f) If you draw the three histograms on the same scale, what do you see? Explain what you observe using the CLT.

3. Justify the F-test using the asymptotic argument rather than assuming the error term to have the normal distribution.

4. Consider the following regression model under i.i.d. sampling:
$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$$
   where x_{3i} and u_i may be correlated but that x_{2i} and (x_{3i}, u_i) are independent.
   (a) Show that the OLS estimator of β_2 when y_i is regressed on the constant term, x_{2i}, and x_{3i} is consistent and asymptotically normal. Denote this estimator by $\hat{\beta}_2$.
   (b) Show that the OLS estimator of β_2 when y_i is regressed on the constant term and x_{2i} is consistent and asymptotically normal. Denote this estimator by $\tilde{\beta}_2$.
   (c) Discuss the asymptotic relative efficiency of $\hat{\beta}_2$ and $\tilde{\beta}_2$.
   (d) Relate this result to the situation where x_{2i} is randomly given to observation i so that it is independent with (x_{3i}, u_i). Since x_{3i} corresponds to observed variable, it may be correlated with unobserved variable.

5. Consider the following regression model with a measurement error problem:
$$y_i = 1 + x_i^* + z_i + u_i$$
   where
$$\begin{pmatrix} x_i^* \\ z_i \\ u_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$$
   and $x_i = x_i^* + v_i$ where $v_i \sim N(0, \sigma_v^2)$ where $\sigma_v^2 = 1$ and $\rho = 0$.

(a) Generate 100 i.i.d. data from the model above and regress y_i on the constant term, x_i and z_i to verify that there is the attenuation bias. State a way you can simulate data to verify the direction of inconsistency proved in class.

(b) What will happen to the size of inconsistency if σ_v^2 increases? Verify this by simulating the data with $\sigma_v^2 = 2$.

(c) What will happen to the size of inconsistency if ρ increases? Verify this by simulating the data with $\rho = 0.5$.

6. Consider the following simultaneous equations model

$$\begin{aligned} y_i^S &= \beta_1^S + \beta_2^S p_i + u_i^S \\ y_i^D &= \beta_1^D + \beta_2^D p_i + u_i^D \end{aligned}$$

where

$$\begin{pmatrix} u_i^S \\ u_i^D \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_S^2 & 0 \\ 0 & \sigma_D^2 \end{pmatrix} \right)$$

and p_i is determined via $y_i^S = y_i^D$. Let $y_i = y_i^S = y_i^D$.

(a) What is the equilibrium p_i in terms of u_i^S and u_i^D? State clearly the conditions under which the solution exists.

(b) Obtain the probability limit of the OLS estimator when observed y_i is regressed on the constant term and p_i under the standard assumptions.

(c) Discuss conditions under which the probability limit of the OLS estimator of the coefficient on p_i equals β_2^S or β_2^D.