Econometrics 522A
Spring, 2019
Hidehiko Ichimura, TA: Zejin Shi

PS #4 (Due 13 Feb, 2019)

1. Study Lalive and Zweimuller (2009) in the syllabus and answer the following questions.
   (a) Write down a linear regression model which corresponds to the (A) Treatment panel of the Contrast column in TABLE II (p.1379).
   (b) Using (A) Treatment panel of the Contrast column in TABLE II (p.1379), test if there is a difference in the parental leave taking behavior between June and July mothers in year 1. How about in year 2? Be specific about the null hypothesis and the alternative hypothesis along with the significance level you used.
   (c) Write down a linear regression model which corresponds to the Base model and the model with Controls in TABLE III (p.1382).
   (d) Explain why we do not expect the coefficients on the July dummy to change very much after including the constant term when we include controls in the base model. The standard error of the July dummy did not change very much, either. Is this expected?
   (e) What does the coefficient on July dummy variable represent? Can you give a causal interpretation?
   (f) What's the reason for running the regression using "Half-window"?

2. Consider a two regressors linear regression model where

$$\mathbf{X}'_{(i)}\mathbf{X}_{(i)} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

   Compute the effect of the ith observation on the predicted value when the regressor is x_i, by comparing $x_i'\hat{\beta}$ and $x_i'\hat{\beta}_{(i)}$. Discuss when the impact of $y_i - x_i'\hat{\beta}_{(i)}$ is large and small given the size of x_i.

3. Using R or Stata and Python,
   (a) write a general program to compute the influence of the ith observation on the OLS estimate.
   (b) Using the linear regression model from PS #2, 14, identify the five most influential data in terms of affecting the OLS estimate of the coefficient on x_{1i}. Theoretically explain what you expect.

4. Construct a sequence of random variables which converges to 0 in probability but does not converge in rth moment for any $r \geq 1$ nor would it converge almost surely. (Hint: modify an example that shows convergence in probability not implying almost sure convergence to make sure that the moment diverges.)

5. Prove that the convergence of a random variables X_n in distribution to a constant value c, implies convergence in probability of X_n to c by showing the following steps:
   (a) What is the CDF that corresponds to a constant c?
   (b) Let $F_n(t) = \Pr(X_n \leq t)$. Write down the condition for the $X_n \xrightarrow{d} c$.
   (c) For any $\epsilon > 0$, show that $F_n(c+\epsilon) \to 1$ and $F_n(c-\epsilon) \to 0$.
   (d) Show that

$$\Pr(|X_n - c| \leq \epsilon) = \Pr(c - \epsilon \leq X_n \leq c + \epsilon) = F_n(c+\epsilon) - F_n(c-\epsilon) + \Pr(X_n = c + \epsilon).$$

   (e) Use the earlier results to show that $X_n \xrightarrow{p} c$.

6. Let A and B be invertible matrices and define the following matrix

$$\begin{pmatrix} A & C \\ B' & D \end{pmatrix},$$

and assume that the matrix is also invertible. Show that

$$\begin{aligned}
\begin{pmatrix} A & C \\ B' & D \end{pmatrix}^{-1} &= \begin{pmatrix} (A - CD^{-1}B')^{-1} & -A^{-1}C(D - B'A^{-1}C)^{-1} \\ -D^{-1}B'(A - CD^{-1}B')^{-1} & (D - B'A^{-1}C)^{-1} \end{pmatrix} \\
&= \begin{pmatrix} A^{-1} + A^{-1}C(D - B'A^{-1}C)^{-1}B'A^{-1} & -A^{-1}C(D - B'A^{-1}C)^{-1} \\ -D^{-1}B'(A - CD^{-1}B')^{-1} & D^{-1} + D^{-1}B'(A - CD^{-1}B')^{-1}CD^{-1} \end{pmatrix},
\end{aligned}$$

holds by answering the steps below:

(a) Compute the left-hand side of

$$\begin{pmatrix} A & C \\ B' & D \end{pmatrix}\begin{pmatrix} W & Y \\ X' & Z \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0' & I \end{pmatrix},$$

and solve for W, X, Y, and Z to show the first equality should hold.

(b) Explain why, for the same matrices,

$$\begin{pmatrix} W & Y \\ X' & Z \end{pmatrix}\begin{pmatrix} A & C \\ B' & D \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0' & I \end{pmatrix},$$

should hold.

(c) Use this relationship and the earlier results to show that the second equality should hold.