Methods of Economic Investigation II (EC403)
Problem Set #3

Oliver Linton

Instructions: Prepare for the week of 13th November.

1. Let

$$y_i = \alpha + \beta x_i + \varepsilon_i, i = 1, 2, \ldots, n,$$

where the usual assumptions hold. However, some data are missing, due to careless data recording. In fact, complete data on y is available, while x is only present for the first $n_1 = n/2$ observations. Let

$$\begin{pmatrix} \widehat{\alpha}_1 \\ \widehat{\beta}_1 \end{pmatrix} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}_1, \mathbf{y}_1 = (y_1, \ldots, y_{n_1})', \mathbf{X}_1 = \begin{pmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_{n_1} \end{pmatrix}',$$

be the OLS estimates of α and β based on the first n_1 observations. Now consider the augmented estimator

$$\begin{pmatrix} \widehat{\alpha}^* \\ \widehat{\beta}^* \end{pmatrix} = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{y}, \mathbf{y} = (y_1, \ldots, y_n)', \mathbf{X}_* = \begin{pmatrix} 1 & . & . & . & . & 1 \\ x_1 & . & x_{n_1} & \overline{x}_1 & . & \overline{x}_1 \end{pmatrix}',$$

where $\overline{x}_1 = n_1^{-1}\sum_{i=1}^{n_1} x_i$ and $\mathbf{X}_*$ is n by 2.

(a) Derive the bias, if any, of the above estimators.

(b) Show that $\text{var}(\widehat{\alpha}^*) \leq \text{var}(\widehat{\alpha}_1)$, while $\text{var}(\widehat{\beta}_1) = \text{var}(\widehat{\beta}^*)$.
Hint: you may find it convenient to rewrite $\widehat{\beta}^*$ in deviation from means form

2. Construct an example of a random sequence X_n and limit random variable X such that $X_n \xrightarrow{D} X$ but X_n does not converge in probability to X. Suppose that $T(\widehat{\theta} - 2\pi) \xrightarrow{d} N(0, 1)$, then show that $T \sin \widehat{\theta} \xrightarrow{d} N(0, 1)$.

3. Let

$$y_i = \beta x_i + u_i, \quad i = 1, 2, \ldots, n,$$

where the usual assumptions hold about u_i; in particular, $E(u_i^2) = \sigma^2$. Suppose that the data were grouped into J unequal categories, thus: $Y_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} y_i$,

$Y_2 = \frac{1}{n_2}\sum_{i=n_1+1}^{n_2} y_i, \ldots, Y_J = \frac{1}{n_J}\sum_{i=n_{J-1}+1}^{n} y_i$, and $X_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i$, $X_2 = \frac{1}{n_2}\sum_{i=n_1+1}^{n_2} x_i, \ldots, X_J = \frac{1}{n_J}\sum_{i=n_{J-1}+1}^{n} X_i$. Here, $\sum_{j=1}^{J} n_j = n$. Now suppose that we estimate β by the no intercept OLS regression of Y_j on X_j, i.e. $\widehat{\beta} = \{\sum_{j=1}^{J} X_j^2\}^{-1}\sum_{j=1}^{J} X_j Y_j$. Derive the limiting distribution of $\widehat{\beta}$. Construct a test of the null hypothesis that $\beta = 0$ based on the grouped data.Reverse Regression

- Suppose that we have a linear regression model

$$y = x\beta + u,$$

with $k = 1$. Define the reverse regression estimator [from regressing x on y]

$$\widehat{\beta}_r = \frac{y'y}{x'y} = \frac{1}{\widehat{\beta}_{x|y}}.$$

Derive the asymptotic distribution of $\widehat{\beta}_r$ making clear any assumptions you make.