Econometrics I
Problem Set 2
Oct. 11, 2002
Due: Oct. 18, 2002

1. Consider the following formulations of the two-variable PRF.

$$\begin{aligned} \text{Model I}: Y_i &= \beta_1 + \beta_2 X_i + u_i \\ \text{Model II}: Y_i &= \alpha_1 + \alpha_2 (X_i - \bar{X}) + u_i \end{aligned}$$

   (a) Find the estimators of β_1 and α_1. Are they indetical? Are their variances identical?

   (b) Find the estimators of β_2 and α_2. Are they indetical? Are their variances identical?

2. In the regression $Y_i = \beta_1 + \beta_2 X_i + u_i$, suppose we multiply each X value by a constant, say, 2. Will it change the residuals and fitted value of Y? Explain. What if we add a constant 2 to each X value?

3. Let $x_1, x_2, \cdots, x_n$ be a sample of size n from a distribution with mean μ and variance σ^2. Consider the following point estimator of μ:

$$\begin{aligned} \hat{\mu}_1 &= \bar{x}, \text{ the sample mean} \\ \hat{\mu}_2 &= x_1 \\ \hat{\mu}_3 &= \frac{x_1}{2} + \frac{1}{2(n-1)}(x_2 + x_3 + \cdots + x_n) \end{aligned}$$

   (a) Which of these estimators are unbiased?

   (b) Which of these is the most efficient?

   (c) Which of these are consistent.

4. A randome variable X follows the exponential distribution if it has the following probability density function:

$$\begin{aligned} f(X) &= \frac{1}{\theta} e^{-\frac{X}{\theta}} \text{ for } X > 0 \\ &= 0 \text{ elsewhere} \end{aligned}$$

   where $\theta > 0$ is the parameter of the distribution. Show that the ML estimator of θ is $\hat{\theta} = \frac{\sum X_i}{n}$, where n is the sample size.

5. State with reasons whether the following statements are true, false, or uncertain.

(a) Even though the disturbance term in the CLRM is not normally distributed, the OLS estimators are still unbiased.

(b) In the two-variable PRF, if the slope coefficient β_2 is zero, the intercept β_1 is estimated by the sample mean $\bar{Y}$.

6. Show that the simple regression line of Y against X coincides with the simple regression line of X againist Y if and only if $r^2 = 1$, where r^2 is the sample correlation coefficient between X and Y.

7. Consider the regression model

$$\begin{aligned} Y_i &= \beta_1 + \beta_2 X_i + u_i \\ u_i &\sim IN(0, 1)\ i = 1, 2, \cdots T \end{aligned}$$

where u_i's are independently and normally distributed. Suppose the model refers to semiannual data, but the data available are either (1) annual data, where $\bar{Y}_1 = Y_1 + Y_2$, $\bar{Y}_2 = Y_3 + Y_4, \cdots$, and $\bar{X}_1, \bar{X}_2, \cdots$ defined analogously, or (2) moving average data, where $Y_1^* = \frac{Y_1+Y_2}{2}$, $Y_2^* = \frac{Y_2+Y_3}{2}, \cdots$ and $X_1^*, X_2^*, \cdots$ defined analogously.

(a) What are the properties of the error term in the regression model with each set of data: (Y_i, X_i), $(\bar{Y}_i, \bar{X}_i)$ and (Y_i^*, X_i^*)? (i.e. What are the distributions of the error terms?)

(b) Can you use OLS method to estimate β_2 in the case of annual data.