Econometrics I
Problem Set 3
Nov. 8, 2002
Due: Nov. 22, 2002

1. Consider the regression model

$$y_i = \beta_1 + \beta_2 x_i + u_i$$

where $y_i = Y_i - \bar{Y}$ and $x_i = X_i - \bar{X}$. Show that the regression line must pass through the origin?

2. Let $X_i^* = \frac{X_i - \bar{X}}{S_x}$ and $Y_i^* = \frac{Y_i - \bar{Y}}{S_y}$, where S_x and S_y are the standard deviations of X and Y, respectively, in the sample. Show that in the model

$$Y_i^* = \alpha_1 + \alpha_2 X_i^* + u_i$$

$\hat{\alpha}_1 = 0$ and $\hat{\alpha}_2 = r$, the coefficient of correlation between X and Y.

3. Consider the following models:

$$\ln Y_i^* = \alpha_1 + \alpha_2 \ln X_i^* + u_i^*$$
$$\ln Y_i = \beta_1 + \beta_2 \ln X_i + u_i$$

where $Y_i^* = w_1 Y_i$ and $X_i^* = w_2 X_i$, the w's being constants. Establish the relationships between the two sts of regression coefficients.

4. This is a practice question for running regressions by STATA. The data are individuals reporting positive earnings from the Labor Force Survey in 1980, 1990 and 2000. . There are 7 variables in the data file ps3.dat, they are "survey year," "monthly earning (denoted as Y)," "number of schooling years (S)," "age (AGE)," "married (MAR, 1 is married, 0 is unmarried)," "sex (SEX, 1 is male, 0 is female)," and "government employee (GE, 1 for government employee, 0 otherwise). Answer the following questions.

   (a) How many obervations are there in the data set, how many of them are male? female? What proportion of the sample is female in each of these three years?

   (b) One version of the earnings equation is

$$\ln Y = \beta_0 + \beta_1 S + \beta_2 EX + \beta_3 EX^2 + \beta_4 MAR + u$$

   where EX represents years of working experience which is defined as $Age - S - 6$. Run an OLS regression of this simple model for men and women

seperately. How many percentage points will one's earnings increase if one have one more year of schooling? (This is usually referred as the rate of return for education.) Is it significant, under what significance level?

(c) What is the effect of being married on ln y for men and women respectively?

(d) Creat a set of dummy variables indicating the level of education. Group those with less than or equal to 6 years of schooling as "Primary School," those with $S = 9$ as "Junior High," $S = 12$ as 'High School," $S = 14$ as "Junior College," and $S >= 16$ as "College." Replace S in model (1) with those dummy variables and let "High School" be the omitted group. What is the rate of return of a college education for men and women?

(e) Consider another version of earnings equation,

$$\ln Y = \beta_0 + \beta_1 S + \beta_2 EX + \beta_3 EX^2 + \beta_4 MAR + \beta_5 SEX + \beta_6 GE + u$$

Run regressions for 1980, 1990 and 2000 separately. What are the coefficients of GE for these three years? Are they significantly positive or negative? Are they increasing or decreasing over time?