Problem Set I

Problem 1 For this problem you will have to use the data set nls.asc which are available on the website for the course. There are 930 observations on nine variables in this data set, **luwe** (log weekly wage), **educ** (years of education), **exper** (years of experience), **age** (age in years), **fed** (father's education in years), **med** (mothers education in years), **kww** (a test score), **iq** (an iq score), and **white** (indicator for white).

1. Estimate a linear regression model for log wages on education, experience, and experience squared. Report regression coefficients and standard errors. Also report the R^2 and the estimate of the standard deviation of the random error.

2. Predict the effect on average log earnings of increasing everybody's education level by one year.

   Hint: If the regression model is

   $$\log(\mathrm{wage})_\mathrm{i} = \beta_0 + \beta_1 \times \mathrm{educ_i} + \beta_2 \times \mathrm{exper_i} + \beta_3 \times \mathrm{exper_i^2} + \varepsilon_\mathrm{i}.$$

   then the effect of increasing education level of individual i by one year is

   $$\theta_i = \beta_1 - \beta_2 - \beta_3 \cdot (2 \cdot \mathrm{exper_i} - 1)$$

   because one year additional education implies one year less work experience. The average affect is the average of this.

   After you defined this, compare the task to that in lecture 5 where we consider the partial effect in a quadratic model. This is not necessary to solve this problem, but just a reminder to re-check the answer after you have studied lecture 5.

3. Can you obtain the above effect by running a regression with a redefined set of covariates? How? Hint: redefined means that the new covariates are functions of the regressors in the regression model of the first part of this assignment.

4. Assume that the error term in the regression has a normal distribution. Predict the effect on the average level of earnings of the following policy: increase the level of education for those who currently have education below 12 years of education to 12, and leave the level of education for others unchanged. Hint: Use the formula for the mean of the lognormal distribution.

Problem 2 Prove the partial regression formula in lecture 2. Hint: Write the normal equations in partitioned form and solve first for $\hat{\beta}_2$ as a function of $\hat{\beta}_1$. Substitute this solution and solve for $\hat{\beta}_1$. Show that the partial regression formula still holds if we replace y^* by y, i.e. if we do not 'purge' the dependent variable.