Quantitative Method (I)
Department of Economics
National Taipei University
Fall 2017
Homework 4
(Due in Class on Dec. 27)

1. This question is a R exercise and you need the dataset *data_for_homework_4*. The data include public expenditure on education (EE), gross domestic product (GDP), and population (P) for 34 countries in 1980. It is hypothesized that per capita expenditure on education is linearly related to per capita GDP. That is

$$y_i = \beta_1 + \beta_2 x_{i2} + u_i$$

where $y_i = (\frac{EE_i}{P_i})$ and $x_i = (\frac{GDP_i}{P_i})$. Based on the above, answer the following questions.

   (a) Show the results from the function $lm()$.

   (b) Do Breusch-Pagan test and White test using the funtion *bptest()* from the package **lmtest**, respectively. What are their null hypotheses? Do we reject the null hypotheses at the 5% level?

   (c) Use White's heteroscedasticity-consistent standard errors and show the results from the functions *coeftest(,vcov=)* from the packages **lmtest** and **sandwich**. Compared to the question (a), is there a difference in $\widehat{\beta}_2$? and is there a change in its standard error?

   (d) Suppose that $\sigma_i^2 = \sigma^2 x_i$, do a generalized least squares estimation. Specifically, transform the data as $y_i^* = \frac{y_i}{\sqrt{x_i}}$, $x_{i1}^* = \frac{1}{\sqrt{x_i}}$ and $x_{i2}^* = \frac{x_i}{\sqrt{x_i}}$ first and then show the results from the function $lm()$. Compared to the questions (a) and (c), is the standard error of $\widehat{\beta}_2$ here the smallest? Note: the intercept needs to be removed from the function.

   (e) Obtain the residuals from the question (d) using the function *residuals()* and regress its squared term on x_{i1}^* and x_{i2}^* using the funtion *lm()*. Based on the F test at the 5% level, is there evidence of heteroscedasticity in this regression model? Note: the intercept does not need to be removed from the function.

2. Suppose that you have a sample of size of n on variables y and x. You are interested in the model $y_i = \beta x_i + u_i$, where β is an unknown parameter. u_i follows a normal distribution with zero mean. About the variance, you find heteroscekdasticity in this dataset and hence propose $\sigma_i^2 = (\delta w_i)^2$, where δ is the second unknown parameter and w_i is a potential variable that is able to explain σ_i. Based on the above, derive the maximum likelihood estimators of β and δ. Note: no need for matrix algebra.