Economics 142
Problem Set Number 2

1. Consider a sample of size N of observations on Y_i. Assume that $E[Y_i] = \mu$, and $E[(Y_i - \mu)^2] = \sigma^2$.

a) Consider the sample mean $\overline{Y} = \frac{1}{N}\sum_i Y_i$. Show that

$$E[(\overline{Y} - \mu)^2] = \sigma^2/N.$$

b) Consider the sum:

$$D = \frac{1}{N}\sum_i (Y_i - \overline{Y})^2$$

Using the fact that $(Y_i - \overline{Y}) = ((Y_i - \mu) + (\mu - \overline{Y}))$ show that:

$$D = \frac{1}{N}\sum_i (Y_i - \mu)^2 - \frac{1}{N}\sum_i (\overline{Y} - \mu)^2$$

c) Using your answer in (a) show that $E[D] = \frac{N-1}{N}\sigma^2$.

d) Using (c), show that

$$E[\frac{1}{N-1}\sum_i (Y_i - \overline{Y})^2] = \sigma^2$$

2. Suppose we have a sample of size N made up of two groups, denoted 0 and 1. Let D_i be a dummy variable with $D_i = 1$ indicating membership in group 1. Finally, let N_0 and N_1 represent the size of the two subgroups (so $N = N_0 + N_1$).

a) Consider an OLS regression model:

$$y_i = \alpha + \beta D_i + u_i$$

Using the first order conditons for the OLS estimates $(\widehat{\alpha}, \widehat{\beta})$, show that

$$\begin{aligned} \widehat{\alpha} &= \frac{1}{N_0}\sum_{i \in 0} y_i, \\ \widehat{\alpha} + \widehat{\beta} &= \frac{1}{N_1}\sum_{i \in 1} y_i \end{aligned}$$

b) Now consider a case where there are 3 groups: $0, 1, 2$ and there are 2 dummies: $D_{1i} = 1$ if i is in group 1, and 0 otherwise, and $D_{2i} = 1$ if i is in group 2, and 0 otherwise, with the regression model:

$$y_i = \alpha + \beta_1 D_{1i} + \beta_2 D_{2i} + u$$

i) find the first order conditions for minimizing the sum of squared residuals, which define the OLS estimator.

ii) using the FOC show that the OLS estimators for the 3-group model will have

$$
\begin{aligned}
\widehat{\alpha} &= \frac{1}{N_0}\sum_{i\in 0} y_i \\
\widehat{\alpha}+\widehat{\beta}_1 &= \frac{1}{N_1}\sum_{i\in 1} y_i \\
\widehat{\alpha}+\widehat{\beta}_2 &= \frac{1}{N_2}\sum_{i\in 2} y_i
\end{aligned}
$$

3. In Lecture 3 we showed that the j^{th} row of the population regression coefficient β^* from the model

$$y_i = x_i\beta^* + u_i$$

can be obtained by first getting the residual from an auxilliary regression of x_{ji} on all the other $x's$:

$$x_{ji} = x'_{(\sim j)i}\pi + \xi_i$$

then forming:

$$\beta_j^* = E[\xi_i^2]^{-1}E[\xi_i y_i]$$

Suppose that we also "residualized" the dependent variable using an auxilliary regression of y_i on all the other $x's$:

$$y_i = x'_{(\sim j)i}\lambda + \phi_i$$

Show that its also true that:

$$\beta_j^* = E[\xi_i^2]^{-1}E[\xi_i\phi_i]$$

(In other words, we could residualize BOTH x_{ji} AND y_i and still get the same right anwer).