Assignment 2.

1. Discuss the following statement:
"The Gauss-Markov theorem implies that the Maximum likelihood estimator will always have smaller variance than the Least squares estimator".

2. Discuss the main implications of not knowing the exact distribution of the disturbance ($\boldsymbol{\varepsilon}$) for the Least squares and Maximum likelihood estimators.

3. Show that $\mathbf{b}$ and s^2 are independent random variables.

4. Suppose data are generated by the following model

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i,$$

where x is an independent variable, i is an index for individuals and ε_i are independently (also from all x_i) normally distributed $N\left(0, \sigma^2\right)$. The density function of y_i, for a given x_i, is then given by

$$f\left(y_i \mid \boldsymbol{\beta}, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\frac{\left(y_i - \beta_1 - \beta_2 x_i\right)^2}{\sigma^2}\right].$$

a) Give an expression for the log-likelihood contribution of observation i, $\ell_i\left(\boldsymbol{\beta}, \sigma^2\right)$. Explain why the loglikelihood function of the entire sample is given by

$$\ell\left(\boldsymbol{\beta}, \sigma^2\right) = \sum_{i=1}^{N} \ell_i\left(\boldsymbol{\beta}, \sigma^2\right).$$

b) Obtain the first order conditions and show that they all have expectation zero for the true parameter values.

c) Suppose that x_i is a dummy variable. $x_i = 1$ for males (the first N_1 observations) and $x_i = 0$ for females. Derive the first order conditions for Maximum likelihood and show that the ML estimators are given by

$$\widehat{\beta}_1 = \frac{1}{N - N_1} \sum_{i=N_1+1}^{N} y_i, \quad \widehat{\beta}_2 = \frac{1}{N_1} \sum_{i=1}^{N} y_i - \widehat{\beta}_1.$$

What is the interpretation of the true parameter values β_1 and β_2 in this case?

d) Present two ways of estimating the asymptotic covariance matrix for $\widehat{\boldsymbol{\beta}}$ and compare the results.

5. Let y_i denote the number of times individual i buys tobacco in a given month. Suppose a random sample of N individuals is available, for which we

observe values 0,1,2,...... Let x_i be an observed characteristic of these individuals. If we assume that, for given x_i, y_i has a Poisson distribution with parameter $\lambda_i = \exp\{\beta_1 + \beta_2 x_i\}$, the probability mass function of y_i conditional upon x_i is given by

$$P\{y_i = y \mid x_i\} = \frac{e^{-\lambda_i}\lambda_i^y}{y!}$$

a) write down the loglikelihood function for this so-called Poisson regression model.

b) Derive the score vector. Using the Poisson distribution implies that $E(y_i \mid x_i) = \lambda_i$. Show that the score has expectation zero.

d) Derive an expression for the information matrix. Use this to determine the asymptotic covariance matrix of the ML estimator and a consistent estimator for this matrix.

6. The following regression equation is estimated as a production function for Q :

$$\begin{aligned} \ln Q &= 1.37 + \underset{(0.257)}{0.632} \ln K + \underset{(0.219)}{0.452} \ln L \\ R^2 &= 0.98 \qquad cov(b_K, b_L) = 0.055 \qquad n = 25 \end{aligned}$$

where standard errors are given in parentheses, K is capital and L is labor.

a) Test the hypothesis that capital and labor elasticities of output are identical.

b) Test that there are constant returns to scale.

c) Assume that you could not reject the hypothesis in b). Show how you could use that this hypothesis applies in the population to increase the efficiency of your estimator.