Exercise Set 1

Problem 1

Consider the following linear regression model:

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

where $\beta = (\beta_1, \beta_2)'$ is a vector of unknown parameters, and x_i is a one-dimensional observable variable. We have a sample of $i = 1, ..., N$, independent observations and assume that the error terms ε_i are normally and independently distributed with mean zero and variance σ^2, $[N(0, \sigma^2)]$, independent of all x_i. The density function of y_i (for a given x_i) is then given by

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

1.1 Give an expression for the loglikelihood contribution of observation i, $log\ L_i\ (\beta, \sigma^2)$. Explain why the loglikelihood function of the entire sample is given by

$$log\ L\ (\beta, \sigma^2) = \sum_{i=1}^{N} log\ L_i\ (\beta, \sigma^2).$$

1.2 Calculate the first derivative $\frac{\partial\ log\ L_i(\beta, \sigma^2)}{\partial \beta}$ with respect to β_1 and β_2.

1.3 Suppose that x_i is a dummy variable equal to 1 for males and 0 for females, such that x_i for $i = 1, ..., N_1$ (the first N_1 observations) and $x_i = 0$ for $i = N_1 + 1, ..., N$.

Derive the first order conditions for maximum likelihood estimators for β and show that the estimators are given by:

$$\hat{\beta}_1 = \frac{1}{N - N_1} \sum_{i=N_1+1}^{N} y_i, \qquad \hat{\beta}_2 = \frac{1}{N_1} \sum_{i=1}^{N_1} y_i - \hat{\beta}_1$$

What is the interpretation of these two estimators? What is the interpretation of the true parameter values β_1 and β_2?

1.4 Suppose that we are interested in the hypothesis $H_0 : \beta_2 = 0$ with alternative $H_1 : \beta_2 \neq 0$. Tests can be based upon the Likelihood ratio, Lagrange multiplier or Wald principle. Explain what these three principles are.

1.5 Discuss for each of the three tests what is required to compute them.

Problem 2

This problem illustrates concepts of maximum likelihood estimation in R using the example of tossing a globe in McElreath (2020).

2.1 Create fake data: Simulate realizations of a Bernoulli random variable for a sequence of $n = 100$ trials. Assume that the true proportion in the population is $\pi = 0.7$.

2.2 Defining the log-likelihood: Use the Bernoulli random variable, y_i, write the likelihood function and log-likelihood function. Provide the computer code in R.

2.3 Maximize the likelihood function by using Grid search in R.

2.4 Do the maximization problem by using R base: optim. Compute the variance and standard errors using the inverse of the negative actual Hessian. Compare the result of the numerical optimization to the result of the gridsearch and the analytical result.

2.5 Do maximization by using Newton- or quasi-Newton (package maxLik). Compute the variance and the standard errors using the inverse of the negative actual Hessian.

2.6 Do maximization by using BHHH algorithm. Compute the variance and standard errors using the inverse of the negative actual Hessian.

Problem 3

Assume that we have independent observations of a random variable y that is normally distributed with unknown mean μ and known standard deviation $\sigma = 1$, i.e. $y \sim N(\mu, 1)$. Suppose the goal is to estimate the unknown parameter μ using data on $n = 10$ realizations of the random variable y shown below.

(i)	**(y)**
1	-1.560476
2	-1.230177
3	0.558708
4	-0.929492
5	-0.870712
6	0.715065
7	-0.539084
8	-2.265061
9	-1.686853
10	-1.445662

3.1 Write down the likelihood function and the log likelihood function.

3.2 Derive the score function and solve the first order condition for the maximum likelihood estimator of μ.

3.3 Give the Fisher information and the variance of the maximum likelihood estimator of μ.

3.4 Use the Newton-Raphson algorithm (or a quasi-Newton method) to numerically compute the maximum likelihood estimate of the unknown parameter μ. Report a 90% confidence interval. Use the data in the above table.

Problem 4

The dataset `vote1` in the R package `wooldridge` contains information on election outcomes and campaign expenditures for 173 two-candidate races in the U.S. (candidate A vs. candidate B). Let *voteA* be the percentage of the vote received by candidate A and *shareA* the percentage of total campaign expenditures by candidate A. We will analyze whether spending more relative to one's opponent is associated with a higher vote share.

4.1 Open `vote1` in R and make yourself familiar with the data. Have a look at the scatter plot of vote share and campaign expenditure of candidate A.

4.2 Estimate the model $voteA = \beta_0 + \beta_1 shareA + u$ using OLS. Interpret the coefficient on shareA.

4.3 Now compute the maximum likelihood estimates of the parameters. Compare the results obtained through maximum likelihood estimation with those from ordinary least squares.

References

McElreath, R. (2020). *Statistical rethinking- A Bayesian course with examples in R and Stan*. CRC press

Verbeek, M. (2012). *A Guide to Modern Econometrics*. John Wiley & Sons, Chichester.

Wooldridge, J. (2013). *Introductory Econometrics. A modern approach*.