Econometrics I, ECO341A, Summer 2023
Homework III (100 points)

Instructor: M.A. Rahman

Deadline: 4:00 pm, July 7, 2023.

Please read the instructions carefully and follow them while writing answers.

- *Solutions to homework should be written in A4 size loose sheets. If you are not comfortable writing on white sheets, please ask for biology paper in Tarun Book Store.*
- *Questions should be answered in order as they appear in the homework. Every new question should begin in a new page. Please number all the pages of your homework solution.*
- *Please leave a margin of one inch from top and one inch from left. Staple the sheets on the top-left.*
- *Matlab assignments (if any) and written answers should be together and in order.*

---

1. (10 points) Let $Y_1, \cdots, Y_n$ be a random sample from a $N(\beta, \sigma^2 = 1)$ and suppose we are interested in testing the hypothesis, $H_0 : \beta = \beta_0$ versus $H_1 : \beta \neq \beta_0$. Derive the expression for Wald, LR and LM test statistics. Are they all different? Explain your finding. (Read the class notes and the article that was sent via email).

2. (5+5=10 points): Suppose we observe $X_1, X_2, \ldots, X_n$ independent Bernoulli (p). The typical parameter of interest is p, but another popular parameter is $p/(1-p)$, the *odds*. For example, if the data represent the outcomes of a medical treatment with $p = 2/3$, then a person has odds $2 : 1$ of getting better. Based on this information, answer the following.

(a) Write the log-likelihood function and find the ML estimator for p, denote it as $\hat{p}_{mle}$ or simply $\hat{p}$ for simplicity. Find the variance of $\hat{p}$.

(b) As we would typically estimate the success probability p with the observed success probability $\hat{p}$, we may consider using $\frac{\hat{p}}{1-\hat{p}}$ as an estimate of $g(p) = \frac{p}{1-p}$. Find $V\big[g(\hat{p})\big]$ using the Delta method.

3. (7+8=15 points): Sometimes, it is necessary to combine the non-sample information in a regression with those contained in the sample observations (y, X). If information of this type is available, it may be stated in the form of a set of linear relations or linear equality restrictions $R\beta = r$, where R is a $(J \times k)$ known prior information design matrix of rank $J \leq k$ and r is a $(J \times 1)$ vector of known elements.

The minimization problem is then modified as follows:

$$\begin{aligned} \min_{\arg\beta} S &= (y - X\beta)'(y - X\beta) \qquad \text{(subject to)} \\ & R\beta - r = 0. \end{aligned}$$

The resulting Lagrangian for the minimization problem is,

$$\mathcal{L} = (y - X\beta)'(y - X\beta) - \lambda'(r - R\beta),$$

where λ is a $(J \times 1)$ vector of lagrange multiplier. Based on this setting, answer the following.

(a) Show that the restricted least squares estimator (or maximum likelihood estimator) has the following expression,

$$\hat{\beta}^* = \hat{\beta} + (X'X)^{-1}R'\big[R(X'X)^{-1}R'\big]^{-1}(r - R\hat{\beta}),$$

where $\hat{\beta}^*$ is the restricted least squares estimator (RLSE) and $\hat{\beta}$ is the OLS estimator.

(b) Show that $E(\hat{\beta}^*) = \beta$ and $V(\hat{\beta}^*) = \sigma^2 M^*(X'X)^{-1}M^*$, where

$$M^* = \Big[I_n - (X'X)^{-1}R'\big[R(X'X)^{-1}R'\big]R\Big].$$

4. (4+6+15=25 points). Consider a linear regression model with dependent variable y, data matrix X (including a column of ones), coefficient vector β and error vector ϵ. The sample size is n. Assume the model satisfies all the CLRM assumptions, except for homoscedasticity. Specifically, a fraction of observations αn with $0 < \alpha < 1$ is associated with error terms that have variance σ_1^2, while the remaining $(1 - \alpha)n$ observations have error variance σ_2^2. For convenience, assume your sample is sorted starting with σ_1^2 cases, followed by the σ_2^2 cases. Given the information, answer the following.

(a) Show the explicit contents of the $n \times n$ variance-covariance matrix Ω of the error vector in terms of σ_1^2, σ_2^2, and identity and zero matrices of appropriate dimensions. (Full points will be awarded only if the dimension of each sub-matrix is specified and correct.)

(b) Show the form of the OLS estimator and *derive* its variance, call it $V(\hat{\beta})$. In light of your finding, discuss the implications of ignoring the heteroscedasticity problem and using the conventional expression for variance-covariance of OLS estimator (i.e., $s^2(X'X)^{-1}$) to derive standard errors and t-values.

(c) Now assume that there is a strong indication that the group-wise heteroscedasticity is driven by an observed indicator variable d which takes the value of "0" for all σ_1^2 cases, and "1" for all σ_2^2 cases. Outline in few lines how you would derive a feasible GLS estimator ($\hat{\beta}_{FGLS}$) for this case. Make sure to show the explicit *skedastic* function you would use. How would you proceed if you did not use the *skedastic* function?

(Note: A *skedastic* function links individual variance terms to observed data. For the multiplicative heteroskedastic model, the *skedastic* function can be written as $\log(\sigma_i^2) = z_i'\gamma$.)

5. (6+4+10=20 points) The managing partner of an advertising agency is interested in the possibility of making accurate predictions of monthly billings. Monthly data on amount of billings (y, in thousand of dollars) and number of hours of staff time (x, in thousand hours) for the 20 most recent months is presented in the file 'Advertising.xlsx'. A simple linear regression model is believed to be appropriate, but positively correlated error terms may be present.

(a) Fit a simple linear regression model and report the OLS estimates of β, standard error of β and the residuals.

(b) Plot the residuals against time and explain whether you find any evidence of positive correlation.

(c) Conduct a formal test for positive correlation using $\alpha = 0.05$. State the alternatives, decision rule and conclusion. Is the residual analysis in Part (b) in accord with the test result?

6. (6+8+6 = 20 points). Consider the data in the file `TransportChoiceDataset.xlsx`. The objective is to study individuals choice between automobile and transit for trip to work. The dependent variable `depend` takes the value 1 if automobile is chosen and 0 if transit is chosen. The covariates in the model are `intercept`, `dcost`, `cars`, `dovtt` and `divtt`. A description of these variables is present in the file.

(a) Present the descriptive summary of the variables (i.e., mean and standard deviation for continuous variables and count and percentage for discrete variables) in a table.

(b) Estimate Probit and Logit models by regressing the dependent variable `depend` on `intercept`, `dcost`, `cars`, `dovtt` and `divtt`. Present the regression coefficients and the standard errors in a table. Numbers should be reported to 3 digits after the decimal.

(c) Calculate the sum of the log-likelihood, Akaike Information Criterion, Bayesian Information Criterion and Hit-rate for the Probit and Logit models.

(`Important:  Please bring the output from Q6 to the final examination.`)