Applied Problems 1

Research Methods II
Computational Economics
University of Exeter, 2020

Description:
Please attempt the below two problems. All of the below is written assuming that you will attempt these problems in MATLAB. However, if you wish to attempt these problems in some other matrix based language, this is fine. If you attempt the problems in MATLAB, you will likely be able to use a some of the work that we do together in class on 31/01/2020 to help in responding to the first problem.

Please send your answers to damian.clarke@protonmail.com by Thursday 6, Feb, 2020 at 23:59. Your answers should consist of the MATLAB functions you write, a README file explaining these functions, and a brief document explaining your response to each of the questions below.

The mark will be assigned in the following way: Does the function for question 1 work in the way it should? (45%); Does the function for question 2 work in the way it should? (30%); Clarity of responses in the document submitted (10%); Clarity of documentation to the code including README and comments within the code (15%).

If you have any questions, please write to damian.clarke@protonmail.com. We will discuss this document further, as well as potential ways to attack these problems, during class on 31/01/2020 when we begin working together in MATLAB. It is not expected that you should be able to respond to these problems before we undertake this discussion in class.

Questions:

1. In classes we will discuss *functions*, and an application to a particular example: the Ordinary Least Squares Regression. In this question, you should write a function `OLS.m` which will be based on the example we see in class, as well as a number of extensions.

   (a) Take the function `OLS.m` which we started with in class. Within this function calculate a number of additional statistics: (i) the t-statistic associated with the test that each coefficient in the model is equal to 0 (one per coefficient); (ii) The p-value associated with this statistic, and (iii) the 95% confidence interval for each parameter, which is calculated as $[\widehat{\beta} \pm t_{0.025}^{N-K} \times se(\widehat{\beta})]$, where $\widehat{\beta}$ and $se(\widehat{\beta})$ refer to the point estimate and its standard error, and $t_{0.025}^{N-K}$ the 2.5$^\text{th}$ percentile of the t distribution with $N - K$ degrees of freedom.
   Hint: `tcdf()` may be useful in calculating the p-value, and `tinv()` may be useful in calculating confidence intervals.

   (b) Add a new section to the function so that the function prints all of these elements neatly on the MATLAB output. It may be useful to model this on Stata's regression output (as displayed in class), though you may consider alternative ways to display the information if you prefer.

   (c) Currently `OLS.m` estimates standard errors which permit valid inference in the case of homoscedasticity. Add and *argument* to the function allowing for standard errors which are robust to heteroscedasticity ("White's standard errors"), which are calculated based on the variance-covariance matrix: $V_H = (X'X)^{-1}(X'\text{diag}(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N)X)(X'X)^{-1}$ (the standard errors are the square root of the diagonal of this matrix).
   Hint: If you are comparing your standard errors to standard errors returned by Stata with the robust option they will be slightly different given that Stata applies a finite sample correction (described on pp. 21–22 here).

2. LaTeX is a system for type-setting documents based on a particular syntax. For example, in LaTeX, the following code (when compiled) makes the table displayed below the code:

LaTeX Code:

```
\begin{tabular}{lcc}
  \toprule
      & (1)   & (2)      \\
      &  y    &  y       \\ \midrule
  X 1 & 1.1   & 1.1      \\
      & (0.5) & (0.6)    \\
  X 2 & 2.2   & 2.2      \\
      & (0.7) & (0.8)    \\
  X 3 & 3.3   & 3.3      \\
      & (0.8) & (0.9)    \\ \midrule
  Observations& 74 & 74  \\ \bottomrule
\end{tabular}
```

Output:

	(1)	(2)
	y	y
X 1	1.1	1.1
	(0.5)	(0.6)
X 2	2.2	2.2
	(0.7)	(0.8)
X 3	3.3	3.3
	(0.8)	(0.9)
Observations	74	74

Write a function in MATLAB which takes your regression results from `OLS.m` and automatically writes a text file in LaTeX format that lets you display these regression results, where for each variable, first the coefficient is displayed, and then the standard error is displayed below in parentheses. Note that this function is along the lines of `outreg2` or `estout` in Stata, or `stargazer` in R.

Hint: The fprintf() function in MATLAB may be useful for this task.