University of Victoria **Department of Economics**

ECON 546: Themes in Econometrics
Assignment 3

Due: Monday 15 March, 4:30 p.m.

Question 1:
Suppose that x is a random variable that follows a Pareto distribution, so its p.d.f. is:

$$p(x) = c\,x^{-(c+1)} \;;\quad 1 \leq x < \infty \;;\quad c > 0.$$

The second moment of x is defined only if $c > 2$, so it would be of some interest to test $H_0 : c = 2$.

(a) Construct the Likelihood Ratio Test statistic for testing H_0, against the alternative hypothesis, $H_A : c \neq 2$, assuming that we have a sample of "n" independent observations on x.

(b) Construct the corresponding Wald Test statistic for this problem.

(c) Construct the corresponding Lagrange Multiplier Test statistic for this problem.

(d) Suppose that we have a sample of 200 independent observations from a Pareto distribution, such that their *product* is 1.2204032×10^{39}. Apply each of the above three tests – does each lead to the same conclusion?

12 Marks

Question 2:
Appended to this question sheet is an EViews program that can be used to simulate the large-sample power (and significance level) of the Wald test for testing H_0: $\mu = \mu_0$ *vs*. H_1: $\mu \neq \mu_0$ when the data are sampled randomly from $N[\mu, \sigma^2 = 1]$. For this problem, the MLE for μ is the sample average of the data, and theWald test statistic is just the square of the usual z-test statistic. A copy of this code is also on the server as **S:\Social Sciences\Economics\Econ546\ass3.prg** .

Remember that to run an EViews program you need to first set up an associated EViews workfile, with an appropriate sample and any data that are needed. Use this program ***as a guide*** to help you write a program for the following Monte Carlo experiment.

(a) Write a program to conduct a Monte Carlo experiment to determine the extent of any "size distortion" associated with the three tests in Question 1 when they are applied to a sample of 10 observations, but the *asymptotic* 5% critical value is wrongly used.

[**Hint:** the Pareto random number generator in EViews is very unreliable and generates some extreme values for certain values of the parameter (e.g., when c = 2.) This really messes up the results. So, to generate a Pareto random variable, with parameter c, use the following result: If X is Uniform on (0 , 1) then $Y = (1/X)^{(1/c)}$ is Pareto-distributed with parameter c. So a command of the form: `series y=(1/@runif(0,1))^(1/c)` will generate Pareto-distributed values.]

(b) Run the experiment to determine which test has the least size distortion.

(c) Modify the code and run the experiment to compare the ***powers*** of the three tests under the conditions in part (a) of this question.

(d) Is the comparison in part (c) totally fair, given what you found in part (b)? What might you do to deal with any such unfairness if you were to pursue this work further?

20 marks

Question 3:
Use the data in the Excel workbook, **S:\Social Sciences\Economics\Econ546\Astronauts.xls** to answer this question. ***A description of the variables is given below the data records in the Excel file***.

NASA selected the first group of astronauts (Group 1) in 1959. The group comprised seven military men, who were selected from a pool of 500 candidates who had the required jet aircraft flight experience, the necessary engineering training, and were under 5 feet 11 inches in height! Groups 2 and 3 included civilians. Academic qualifications were taken into account from 1964 onwards, and in 1965 six scientist astronauts were selected from a group of 400 candidates who had a doctorate in the natural sciences, medicine or engineering. The astronauts selected from 1978 onwards trained as space shuttle crews.

The data in the file relate only to astronauts who had been, or were, members of the U.S. space program, as at July 2003. A small number of them were born outside the U.S.A. but entered the regular American Space Program. The data ***exclude*** Russian cosmonauts and astronauts from other space agencies (*e.g.*, the European Space Agency) who have trained and flown joint missions with NASA astronauts in recent years.

(a) Load the data into an EViews workfile, and view the "Histogram and Stats" for the variable "FLIGHTS". Explain why it seems sensible to model this series using a Poisson process.

(b) Estimate a ***Poisson Regression*** model that explains "FLIGHTS" as a function of variables that you think may be important. Include a constant in the model, and be careful in your use of the various dummy variables so as to avoid "perfect multicollinearity". Experiment with the model specification, using the Huber/White standard errors, and try and find a version that includes only significant explanatory variables.

(c) Does a military background seem to be a significant factor in determining the number of space flights an astronaut makes (made)? What about gender, year of birth or academic achievement?

(d) Estimate the Poisson regression model with just a constant and "SELECTED" as the explanatory variables and view the graph for the "Actual, Fitted, Residual". In what particular sense does this model seem to be inadequate? Compare this graph with the corresponding one for the Poisson model with "CSMH", "MOONC", "DECEASED" and "EVA" as ***additional*** explanatory variables. Is there any improvement in the graph?

(e) Using this last model, use the Wald test to test if the Poisson model is preferred to the Negative Binomial model. What do you conclude?

(f) Using either the Poisson or Negative Binomial model (as determined by your result in part (e) above), calculate the marginal effects (on the number of flights) of an additional hour of extra-vehicular activity, and of being awarded the Congressional Space Model of Honor.

18 marks

Total Marks = 50

```
' INITIALIZE VARIOUS VALUES
' ------------------------------------------
scalar nrep=1000
scalar mu0=0
scalar sum=0
rndseed 123456
' CALCULATE 5% CRITICAL VALUE FOR CHI-SQUARE WITH v = 1
'-----------------------------------------------------
scalar crit=@qchisq(0.95,1)
!n=500
smpl 1 !n
' START THE MONTE CARLO LOOP
'-------------------------------------------------
for !i=1 to nrep
series y=@rnorm
scalar ybar=@mean(y)
scalar z=(ybar-mu0)/(1/@sqrt(!n))
scalar wald=z^2
if wald>crit then
scalar sum=sum+1
endif
next
'END OF MONTE CARLO EXPERIMENT
'---------------------------------
scalar power=sum/nrep
```