Session 4: Computer-based Exercise

These exercises follow on from Session 3 exercises. This session we will construct likelihood ratio confidence intervals. Recall that π, the proportion parameter, is referred to as `pi`. Before you start, check that you have given Stata a sensible working directory, and start a log file.

Question 1

(a) Recall Question 1 from Session 3 Practical: suppose we observe 3 deaths among 10 subjects. Assuming a binomial model, write down the likelihood for π: Use the following commands to graph the likelihood:

```
clear
range pi 0.01 0.99 99
gen L=pi^3*(1-pi)^7
twoway line L pi
```

What is the likelihood ratio comparing $\pi = 0.1$ with $\pi = 0.5$? [“Likelihood ratio” here is in the simple sense of the ratio of *these* two likelihoods, rather than the technical sense in the lecture, of ratio of a likelihood over its maximum.]

(b) Below are some Stata commands which will allow you to calculate the likelihood ratio (in the technical sense). First we need the maximum value of the likelihood:

```
egen Lmax=max(L)
```

[That finds the maximum of L, and stores it in the column Lmax (each row with the same value)]

Check that this is the value of the variable L corresponding to pi=0.3, using `list` (eg `list if L==Lmax`). Then we plot the likelihood ratio (the maximum of which should be 1) against π, and then the log-likelihood ratio (the maximum of which should be 0):

```
gen LR=L/Lmax
twoway line LR pi
```

[Notice the likelihood ratio is the same shape as the likelihood, with the vertical axis rescaled.]

```
gen logLR=log(LR)
twoway line logLR pi
```

[Plots the log likelihood ratio]

```
twoway line logLR pi if logLR>-4
```

[That restricts the plot to `logLR` values greater than -4]

[The following single command goes over two lines: hence the “/* */”. Don’t type these symbols if entering into the Stata command line. In a do file, however, they will tell Stata to run on the command.]

```
twoway line logLR pi if logLR>-4,yline(-1.92) yaxis(1 2) /*
*/ylabel(-1.92,  angle(horizontal) axis(2))
```

[That command adds a suitably labelled horizontal line at -1.92 to show the “supported range”. The “supported range” is a Bayesian term which in many contexts is numerically equivalent to the classical confidence interval.]

(c) From the above graph (ie visually estimating) what are the approximate 95% likelihood ratio confidence limits for π? Check your accuracy by listing the values of `pi` and `logLR`.

[PTO]

Question 2

(a) Now you will attempt to superimpose on your previous graph the log-likelihood ratio for π based on 30 deaths observed among 100 subjects. Here are some hints:

It is easiest to calculate the log-likelihood directly:

```
gen logL_30=30*log(pi)+70*log(1-pi)
```

Then you can calculate the log-likelihood ratio (call it `logLR_30`) by *subtracting* the maximum value of the log-likelihood (which you will have to obtain using `egen`).

To plot a clear graph it is best to restrict the range of the log-likelihood ratios to values above -4. You can superimpose plots of both log-likelihood ratios like this:

```
twoway (line logLR pi if logLR>-4, yaxis(1 2)) /*
*/(line logLR_30 pi if logLR_30>-4, yaxis(1 2))/*
*/, yline(-1.92) ylabel(-1.92,  angle(horizontal) axis(2))
```

> The ability to superimpose plots flexibly is a useful feature of Stata: note that each line graph has its options (after their comma in the brackets), and then the comma at the beginning of the last line indicates the start of general options that apply to both plots. As before, don't type "/*" etc if entering directly into the Stata command line: but use these if using a do file.

(b) Compare the log-likelihood ratio curves in terms of the maximum likelihood estimates, 'tightness' of curvature, and symmetry.

Question 3

Find the 95% likelihood ratio confidence interval for λ, as given in Question 3 from Session 3 practical. Here is the question again, for your convenience (you may want to re-use your Stata commands from that session).

"In a follow-up study, 8 deaths are observed during 160 person-years. Assuming a Poisson model with rate parameter λ, plot the likelihood curve. You should use 91 values of λ between 0.010 and 0.100 at intervals of 0.001. Ignore constant terms in calculating the likelihood."