Session 3: Computer-based Exercise

In these exercises you will use Stata to inspect some likelihood and log-likelihood curves. Stata commands are given in this font: `stata command`; other statements are comments etc.. and should of course not be typed into Stata. In the Stata commands, π, the proportion parameter, is referred to as `pi`.

Before you start, check that you have given Stata a sensible working directory, and then start a log file with the command `log using filename.log`. The .log extension saves the log file as a text file which you can view in, eg, Word. Note the difference between the numeral `1` and the letter `l` in the Stata font.

Question 1

Suppose we observe 3 deaths among 10 subjects. Assuming a binomial model, find the likelihood for different values of π between 0.1 and 0.9, using the following commands:

```
clear
```

This clears memory.

```
range pi 0.1 0.9 9
```

Variable π is generated to have the 9 values 0.1,0.2,...,0.9. Type `help range` to see how the this works, and also to get in the habit of using Stata help!

Next:

```
gen L=pi^3*(1-pi)^7
```

[L is the likelihood, ignoring constants such as $\binom{10}{3}$ that don't depend on π; such constants do not affect the likelihood shape, and hence do not influence the value of $\hat{\pi}$.

```
list pi L
```

This lists values on screen.

What value of π gives the maximum likelihood? Which of these two values of π, 0.1 or 0.5, is better supported by the data?

Question 2

To make a more effective graph of the likelihood, it is better to use a finer grid for π. Using π = 0.01 to 0.99 in intervals of 0.01, we can draw a graph of the likelihood curve using the following commands:

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

This plots a line graphing *L* against π.

Stata is case-sensitive, so a variable called `L` is different from a variable called `l`.

A more elegant way of doing the graph in Stata:

```
twoway function y=x^3*(1-x)^7,range(0.0 0.99) n(99) ytitle(L) xtitle(pi)
```

(The twoway function command only allows the use of variables y and x, hence the need to retitle in this example.) However, this command only draws the graph: it doesn't leave any variables behind to play with.

Question 3

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 could use 91 values of λ between 0.010 and 0.100 (this gives intervals of 0.001). Ignore multiplicative constant terms in calculating the likelihood. Find the maximum likelihood estimate.

Then plot the log-likelihood curve and verify that the maximum is at the same point.

You could of course use the twoway function command to plot these graphs if you wish.