# ISE 3293/5013 Laboratory 10: Maximum likelihood estimates using Grid and N-R

In this lab we will learn how to use R to make maximum likelihood estimates from a sample. A number of functions will be used and should be well understood. While you are not expected to make your own functions I will expect you to understand the lines of code and be capable of adjusting the function to make it do whatever is required.

Some likelihood functions can be easily manipulated so that analytical solutions are available. However in some statistical applications a likelihood might be made from the contingencies of a practical problem that will not easily render an analytical solution. In such cases numerical methods like what we shall perform in this lab will need to be invoked. The likelihood proceeds directly from the joint distribution where instead of being viewed as a function in the data $\boldsymbol{x}$ it is treated as a function in the parameter $\theta$.

$$f(\boldsymbol{x}|\theta) \to L(\theta|\boldsymbol{x}) \to l(\theta|\boldsymbol{x})$$

The process involves differentiating the log likelihood $l(\theta|\boldsymbol{x})$ and finding the maximum. Though this is the analytical process it should be remembered that all we want is the value of $\theta$ that maximizes the likelihood. We will therefore be able to use a simpler method by making R evaluate the likelihood with a range of parameter values and then find the parameter value corresponding to the maximum. This method is called Grid approximation.

The following is a function that will make maximum likelihood estimates:

```
mymaxlik=function(lfun,x,param,...){
# how many param values are there?
np=length(param)
# outer -- notice the order, x then param
```

```
# this produces a matrix – try outer(1:4,5:10,function(x,y)
paste(x,y,sep=" "))   to understand
```

**z=outer(x,param,lfun) # A**

The outer function uses lfun as operator between the first parameter “x” and the second parameter “param”. Inputs are in vector forms and the output is a matrix structure with elements of “z(i,j) = x(i)◯param(j)”. Operator ◯ is defined by lfun.

```
# z is a matrix where each x,param is replaced with the
function evaluated at those values

y=apply(z,2,sum)


# y is a vector made up of the column sums

# Each y is the log lik for a new parameter value

plot(param,y,col="Blue",type="l",lwd=2,...)

# which gives the index for the value of y >= max.

# there could be a max between two values of the parameter,
therefore 2 indices

# the first max will take the larger indice
```

**i=max(which(y==max(y))) # B**

Finds the maximum value of y.

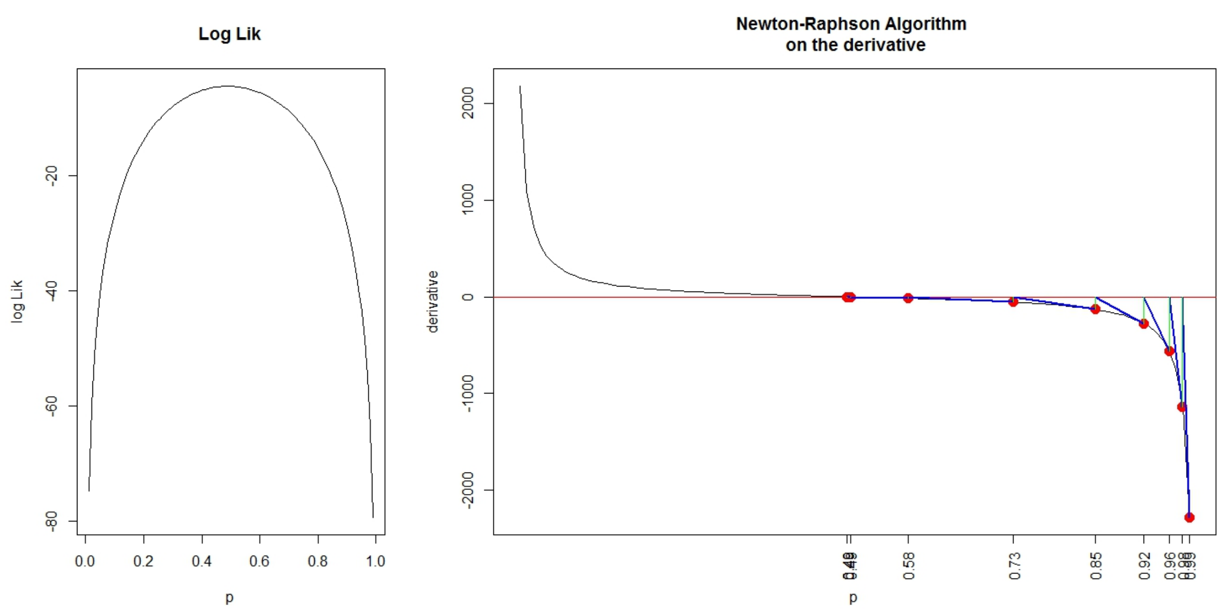
```
abline(v=param[i],lwd=2,col="Red")


# plots a nice point where the max lik is

points(param[i],y[i],pch=19,cex=1.5,col="Black")

axis(3,param[i],round(param[i],2))
```

```
	#check slopes. If it is a max the slope should change sign from + to
	# We should get three + and two -vs
	ifelse(i-3>=1 & i+2<=np, slope<-(y[(i-2):(i+2)]-y[(i-3):(i+1)])/(param[(i-2):(i+2)]-param[(i-3):(i+1)]),slope<-"NA")
	return(list(i=i,parami=param[i],yi=y[i],slope=slope))
}
```

The above code forms the basis of other functions used in this lab.
There is a more sophisticated method available – it uses the Newton Raphson algorithm. Here is a brief description of the method and a function that implements the technique.
First we must identify the essential issue it hand. We need to find where a derivative is zero. This is really just the problem of finding where a function cuts the x axis. We need $x$ such that $f(x) = 0$. The following code gives the recursive algorithm that accomplishes this:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

There is a function that accomplishes this – it is called `myNRML()` – it will produce graphical output similar to the plot below.
The trick to see here is that the function we are attempting to zero is the derivative of the log likelihood. In the case below the derivative is $l'(p)$, that is we need $\hat{p}$ such that $l'(\hat{p}) = 0$, we have theory that handles $f(x) = 0$ so here $l'(p) = f(x)$. The graph on the left is the log lik – you can see that it has a max about p=0.5, and that its slope goes from positive to zero to negative. On the right you can see the derivative $l'(p)$ and it has a zero as calculated with the NR algorithm to be 0.5.



***Tasks***

All output made please copy and paste into **this word file**. Save and place in the dropbox when completed. Anything you are asked to make should be recorded under the question in this document. There will be two files you need to upload:

- a pdf of this document (pdf) or the word file (docx)
- a text file of all the code you used to create answers (txt)

**Note: All plots you are asked to make should be recorded in this document.**

**You are expected to adjust the functions as needed to answer the questions within the tasks below.**

- Task 1
  - Make a folder LAB10
  - Download the file "lab10.r"
  - Place this file with the others in LAB10.
  - Start Rstudio
  - Open "lab10.r" from within Rstudio.
  - Go to the "session" menu within Rstudio and "set working directory" to where the source files are located.
  - Issue the function `getwd()` and copy the output here.

    ```
    "F:/Google Drive - Saied/Courses/02 OU/11 Fundamentals of Engine
    ering Statistical Analysis/02 Labs/10 Lab 10"
    ```

- Task 2
  - Create your own R file and record the R code you used to complete the lab.
  - In the above code for `mymaxlik()`there are two lines marked A and B. Using any resources available explain what each line does
    - Line A (Use the command in the comments to explain what `outer()` does)

      `The outer function uses lfun as operator between the first parameter “x” and the second parameter “param”. Inputs are in vector forms and the output is a matrix structure with elements of “z(i,j) = x(i)◯param(j)”. Operator ◯ is defined by lfun.`
    - Line B

      `Finds the maximum value of y which is the answer.`
  - Suppose that a series of 8 binomial experiments were performed, with each having 20 trials and each with the same probability of success $p$. Find the maximum likelihood estimate for $p$ by first answering the following
    - What is the formula (mathematical – no R code) for the likelihood?

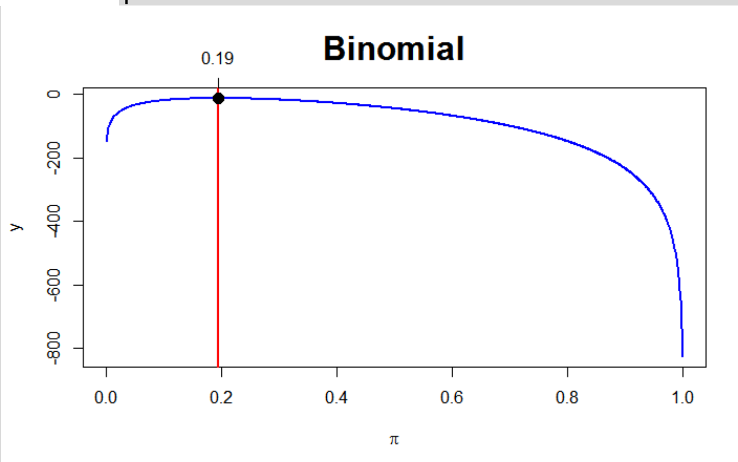
$$\prod \binom{20}{y} p^y (1-p)^{20-y}$$

    - What is the R code for the above likelihood?

$$\prod dbinom(y, 20, p)$$

- Find the maximum likelihood estimate when the following is the data $y = 3,3,4,3,4,5,5,4$ (record the plot).

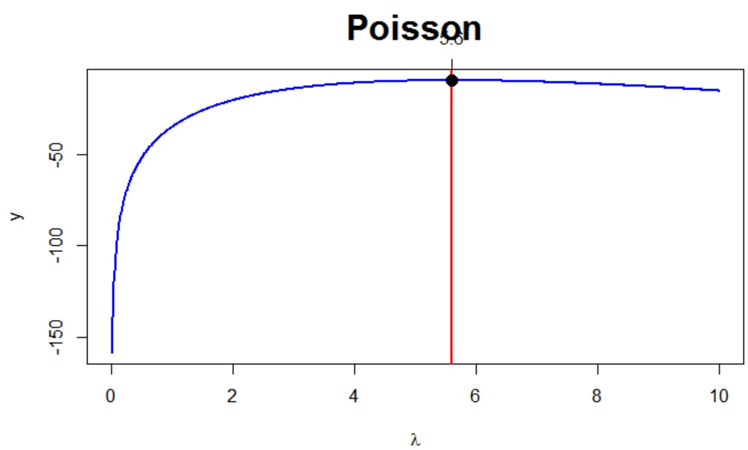
```
p = 0.1919192
```




- Task 3
  - A radioactive source produces protons at an average rate of $\lambda\, protons/sec$. What is the maximum likelihood estimate of $\lambda$ (using graphical methods) when the data collected from a random sample is $y = 4,6,7,6,5$ protons per second?

```
5.595596
```

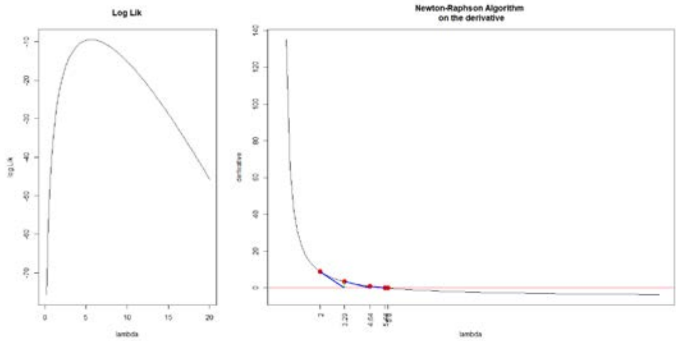
  - Show the graphical solution using `mymaxlik()`




  - What is the algebraic expression for the log likelihood?
  - You may need $p(y) = \frac{e^{-\lambda}\lambda^{y}}{y!}$

$$L(\lambda) = \frac{e^{-\lambda}\lambda^{4}}{4!} . \frac{e^{-\lambda}\lambda^{6}}{6!} . \frac{e^{-\lambda}\lambda^{7}}{7!} . \frac{e^{-\lambda}\lambda^{6}}{6!} . \frac{e^{-\lambda}\lambda^{5}}{5!} = \frac{e^{-5\lambda}\lambda^{28}}{4!\,6!\,7!\,6!\,5!}$$
$$l(\lambda) = log\big(L(\lambda)\big) = -5\lambda + 28log(\lambda) - log(4!\,6!\,7!\,6!\,5!)$$

  - Now use myNRML() to find the maximum likelihood solution, record the following
    - The graphical output




- The command line output

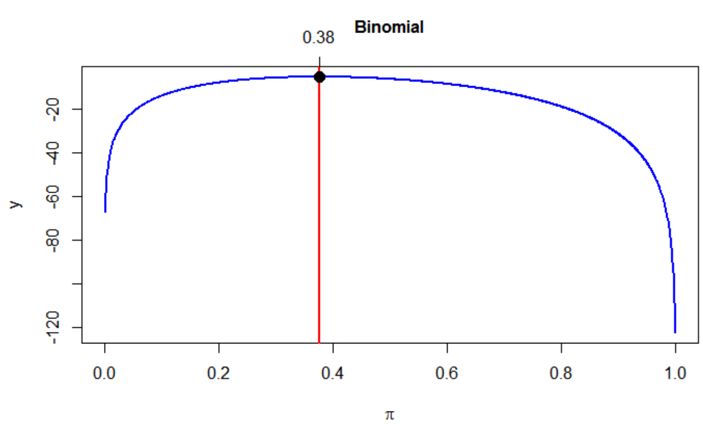
```
$x
[1] 2.000000 3.287233 4.644564 5.436662 5.595322 5.59999
1 5.599999

$y
[1] 8.999996e+00 3.517802e+00 1.028552e+00 1.502188e-01
4.180269e-03 7.839063e-06 1.953993e-08
```

- What is the value of $\hat{\lambda}$ as given by the function `myNRML()`

```
5.599999
```

- Task 4
    - Suppose we have two different binomial experiments using the same biased coin. (I.E The probability of a head is the same)
    - For the first experiment the number of trials was 6 and the number of heads 2, for the second experiment the number of trials was 10 and the number of heads 4.
    - Suppose that $p = probability\ of\ a\ head$. Use `mymaxlikg()` to find the graphical max. lik solution for $p$.

```
x = c(2, 4); n = c(6, 10); param = seq(0, 1, length =
1000);
logbin=function(x,param) log(dbinom(x[1], prob = param,
size = n[1])*dbinom(x[2], prob = param, size = n[2]))
mymaxlik(x,   param   ,lfun   =   logbin      ,xlab   =
expression(pi), main = "Binomial", cex.main = 1)
```




- Task 5

- Suppose an experiment results in a joint density that is the product of a poisson and a binomial.
- That is $p(y_1, y_2|\theta_1\theta_2) = bin(y_1|\theta_1)poiss(y_2|\theta_2)$
- Write down the algebraic expression for $l(\theta_1, \theta_2)$ the log likelihood.

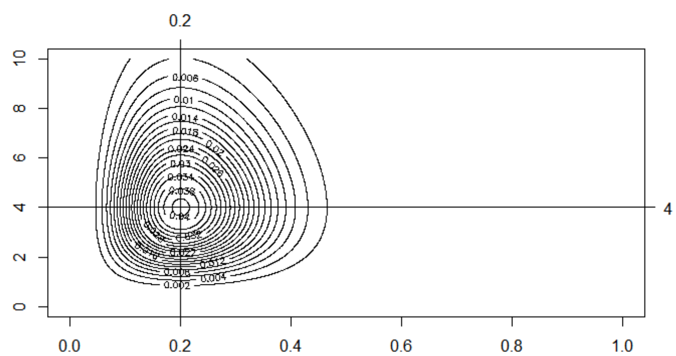
$$l(\theta_1,\theta_2) = log\big(L(\theta_1,\theta_2)\big) = log\left(\binom{n}{y_1}{\theta_1}^{y_1}(1-\theta_1)^{n-y_1}\frac{e^{-\theta_2}{\theta_2}^{y_2}}{y_2!}\right)$$

- Suppose the poisson process has $y_2 = 4$ and the binomial process $n = 20, y_1 = 4$

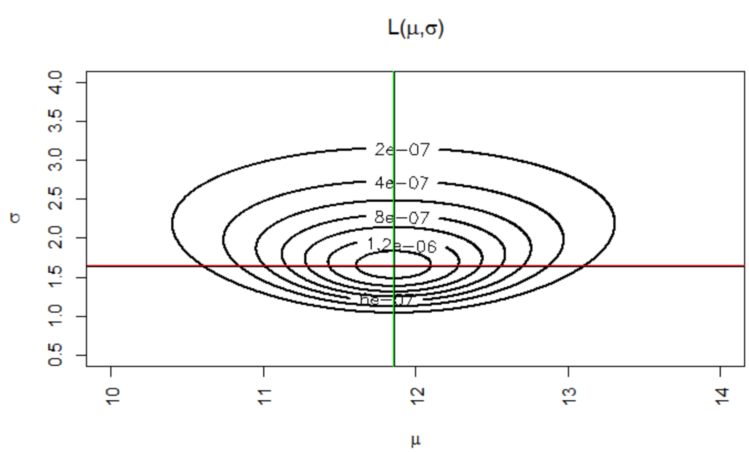
$$l(\theta_1,\theta_2) = log\big(L(\theta_1,\theta_2)\big) = log\left(\binom{20}{4}{\theta_1}^{4}(1-\theta_1)^{20-4}\frac{e^{-\theta_2}{\theta_2}^{4}}{4!}\right)$$

$$l(\theta_1,\theta_2) = log\left(\frac{4845}{24}\theta_1^4(1-\theta_1)^{16}e^{-\theta_2}\theta_2^4\right)$$

- Use `maxlikg2()` to give the graphical max lik solutions for $\theta_1$ $and$ $\theta_2$.

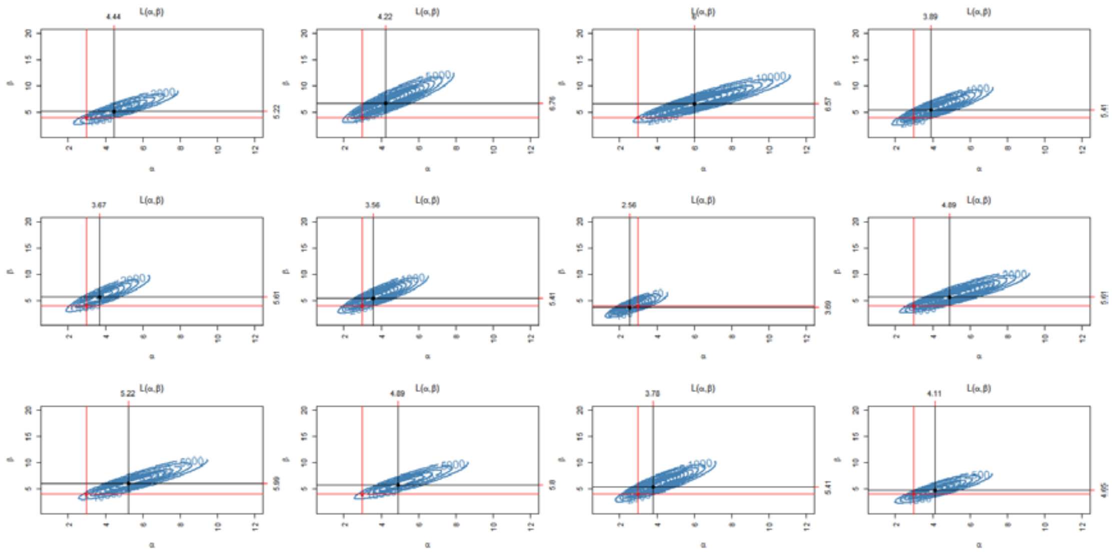



- Task 6
  - Suppose that a normal experiment is performed and scientists are interested in finding maximum likelihood estimates for the population mean and standard deviation.
  - Use the function `mymlnorm()` to produce a graphical solution for $\hat{\mu}$ and $\hat{\sigma}$ when the data are $y = 10,12,13,15,12,11,10$ assuming that $Y_i \sim N(\mu, \sigma)$.

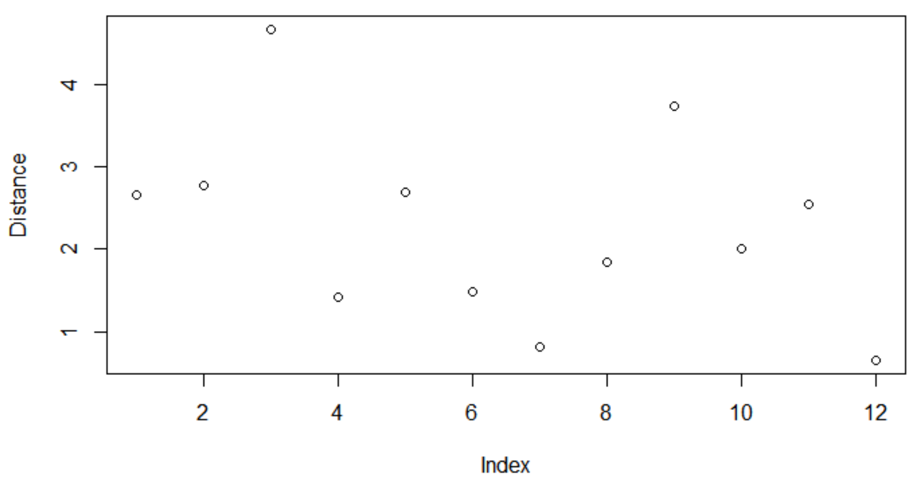



- Task 7

- Suppose that another experiment is performed this time data is generated from a beta distribution. The data is generated as below.
- `sam= rbeta(30,shape1=3,shape2=4)# we know the pop params`
- Using any code available resample from the sample (sam) and create max lik estimates for shape1 and shape2. Hint: You might need the function `mymlbeta()`
- Make 12 plots on the same graphical device.



- Plot the distances between the estimates and the known population values.




################### LAB FINISHES HERE ###############################

- Task 9: Extra for experts!
  - Make a function that will find max lik estimates for a gamma distribution.

```
mymlgamma = function(x, shape , scale,...){  #x sample vector
  na = length(shape) # number of values in alpha
  nb = length(scale)
  n = length(x) # sample size
  zz = c()    ## initialize a new vector
  lfun = function(x, a, b) log(dgamma(x, shape = a, scale = b))   # log lik for gamma
  for(j in 1:nb){
    z = outer(x, shape, lfun, b = scale[j]) # z a matrix
    y = apply(z, 2, sum)
   zz = cbind(zz, y)
  }

  maxl = max(exp(zz))    # max lik
  coord = which(exp(zz)==maxl,arr.ind=TRUE)  # find the co-ords of the max
  aest = shape[coord[1]] # mxlik estimate of shape
  best = scale[coord[2]]

contour(shape,scale,exp(zz),las=3,xlab=expression(shape),ylab=expression(scale),axes=
TRUE, main=expression(paste("L(",shape,",",scale,")",sep="")),...)

  abline(v = aest, h = best)
  points(aest, best, pch=19)
  axis(4, best, round(best,2), col="Red")
  axis(3, aest, round(aest,2), col="Red")
  return(list(x = x, coord = coord, maxl = maxl,maxshape = aest, maxscale = best))
}
```