Maximum Likelihood Estimation and Model Fitting, CMEE MSc. Tin-Yu Hui

Practical 4 (Thursday)

<u>Question 1</u>

    i.   Q2 (iii) from Practical 3. Use R to find the 95% confidence interval for p.

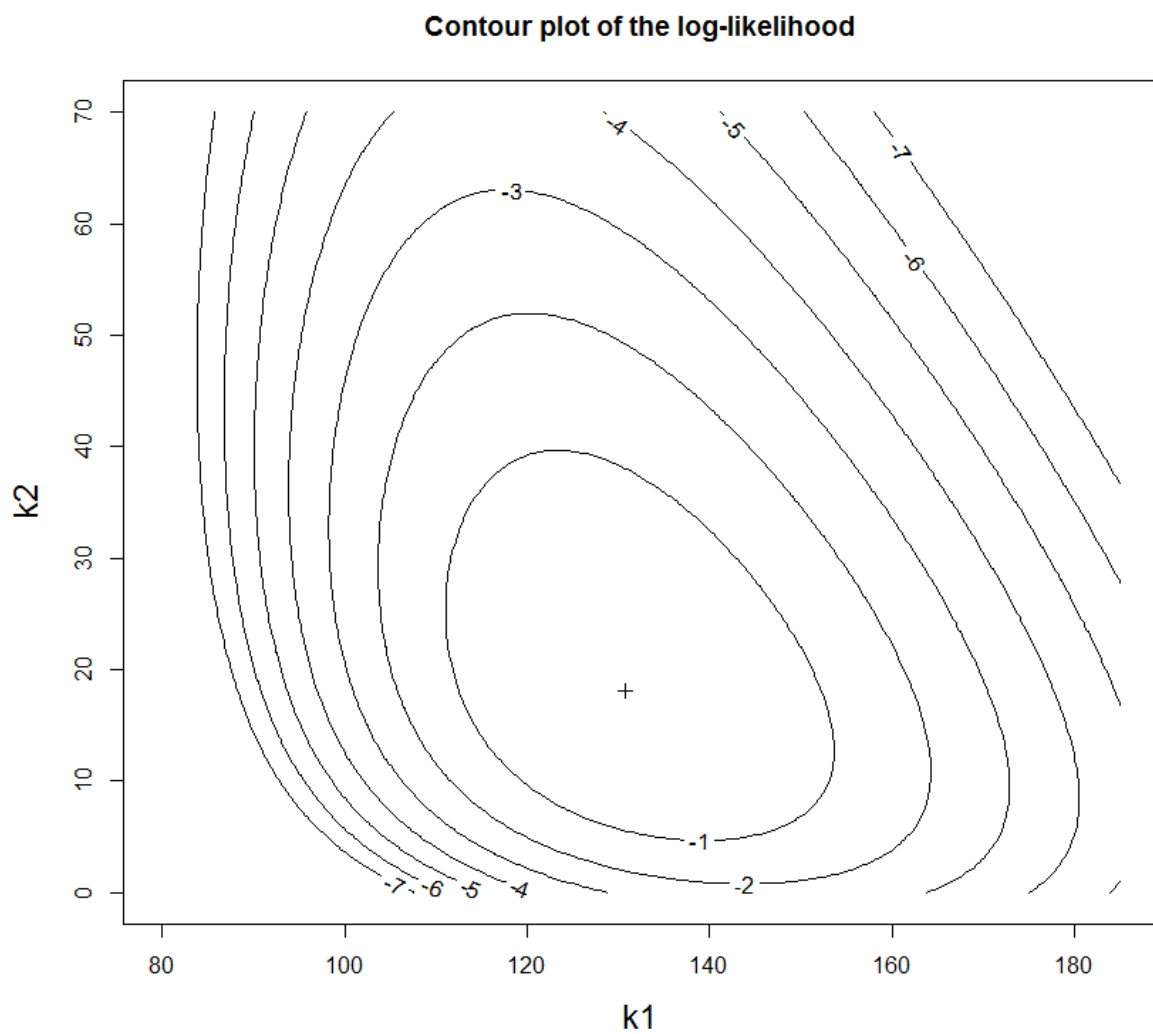   ii.   Find the 95% C.I. for p via normal approximation (by hand).

Question 2

You are given a contour plot of a log-likelihood surface. Can you find the approximate 95%
confidence interval for parameters k1 and k2 respectively by looking at the log-likelihood function?
Also, can you circle the joint 95% confidence region for k1 and k2?

Contour plot of the log-likelihood

Yesterday we fitted several logistic regressions to relate the `State` of a plant to the number of `Flowers` it produced the year before and its `Root` size. We also ran a Likelihood-Ratio Test to show that the interaction between the two explanatory variables is significant. We continue our analysis to find the 95% C.I. for b, the coefficient of `Flowers`. We can no longer visualise the likelihood surface as a contour plot in this four-dimensional parameter space. We have to use profiling instead.

I hope you have already maximised `logistic.log.likelihood.int` and stored the output as `M2`. `M2$par` are the four ML estimates, and `M2$value` is the maximised log-likelihood value.

The ML estimate for b is $\hat{b} = -0.07889$. If we move b away from $\hat{b}$, the log-likelihood drops. The next sensible question will be "what is the log-likelihood value when $b = -0.05$, or when $b = -0.03$?" and so on. A naïve (incorrect!) approach is to evaluate the log-likelihood at $b = -0.03$ while keeping all other estimates unchanged:

```
# THE WRONG APPROACH. SET b=-0.03, AND KEEP OTHERS AT THEIR MLE

logistic.log.likelihood.int(parm=c(M2$par[1], -0.03, M2$par[3],
M2$par[4]), dat=flowering)
```

This is wrong because there is no guarantee that at $b = -0.03$ the log-likelihood is also maximised at $(\hat{a}, \hat{c}, \hat{d})$. The correct procedure is to fix $b = -0.03$, and then re-maximise the log-likelihood with respect to the remaining three parameters. The maximised log-likelihood value under this restricted parameter space is the profile log-likelihood value for b at $b = -0.03$. We can translate this procedure into a function:

```
# PROFILE LOG-LIKELIHOOD FOR b. OBVIOUSLY IT IS A FUNCTION OF b

profile.log.likelihood<-function(b)

{

    f<-function(parm_acd)

    {logistic.log.likelihood.int(c(parm_acd[1], b,

    parm_acd[2], parm_acd[3]), dat=flowering)}


    temp<-optim(c(0, 0, 0), f, control=list(fnscale=-1))

    return(temp$value)

}
```

The argument `parm_acd` for the inner function `f` is a vector of the remaining three parameters `a`, `c`, and `d`. `f` is maximised with respect to these three parameters.

```
# PROFILE LOG-LIKELIHOOD VALUE FOR b AT b=-0.03

profile.log.likelihood(b=-0.03)
```

Now we can profile b in the neighbourhood of $\hat{b}$. Note that each point on the profile log-likelihood involves calling an `optim()`, which can be computationally intensive.

```
# PLOT THE PROFILE LOG-LIKELIHOOD FOR A RANGE OF b, USUALLY AROUND
ITS MLE

b<-seq(-0.19, -0.004, 0.002)

profile.log.likelihood.value<-rep(NA, length(b))

for (i in 1:length(b))

{profile.log.likelihood.value[i]<-profile.log.likelihood(b[i])}


# PLOT THE PROFILE LOG-LIKELIHOOD

plot(b, profile.log.likelihood.value, type='l')

# DRAW A HORIZONTAL LINE WHICH IS 1.92 UNITS BELOW THE MAXIMUM

abline(h=M2$value-1.92, col='red', lty=2)
```

The 95% C.I. for b alone is the range b whose profile log-likelihood drops by no more than 1.92 units (indicated by the red dotted line) from the maximum. Profiling can be applied to any subset of parameters. For example, if we are profiling a and b, then at each given (a_0, b_0), a partial maximisation is performed on the remaining two parameters. And the joint 95% confidence region for (a, b) will be the region where the profile log-likelihood decreases by no more than half of $\chi^2_{0.95, df=2}$ units from the maximum.

[Optional task #1] Find the 95% C.I. for b under normality.

[Optional task #2] Perform the same analysis with `glm()`.