DISCUSSION 4

Suppose the whole population is a mixture of three different subpopulations: African American, Asian, European with mixture proportions as 20 percent, 5 percent, and 75 percent. Also suppose that the divorce probability in African American is two times as high as in European; the divorce probability in Asian is one third as that in European. By randomly selecting 500 persons from the mixed population, 200 of them are divorced.

1. Let \( p \) be the divorced probability in mixed population and \( \beta \) the divorce probability in European population. Show that \( p = 1.17\beta \).

2. Write the formula for the likelihood of \( \beta \) using binomial formula. Calculate the likelihood and the log-likelihood of the divorce probability as 0.5 among Europeans (\( \beta = 0.5 \)).

3. Find the maximum likelihood of \( \beta \). This can be done in several steps.
   (a) Make a fine sequence of \( \beta \) from 0.05 to 0.85.
   (b) Evaluate the likelihood for each of these values. Plot the log-likelihood curve, the log-likelihood values versus the \( \beta \) values.
   (c) Find, the \( \beta \) value in your list that has the largest log-likelihood.

4. Use the likelihood ratio test to test the null hypothesis \( \beta = 0.5 \). Indicate the value of \( X^2 = 2(\log(\hat{\beta}) - \log(\beta_0)) \), the null distribution for this test, the p-value, and the final conclusion.

5. Determine the critical value \( C \) for the likelihood ratio test above, such that \( X^2 > C \) leads to a p-value \( p < 0.05 \).

6. Calculate the \( X^2 \) statistic for testing each \( \beta \) value in the list. Then determine a 95 percent confidence interval for \( \beta \) by finding the limits \( (\beta_1, \beta_2) \) such that \( X^2 < C \) for all beta between \( \beta_1 \) and \( \beta_2 \) and \( X^2 > C \) otherwise.
Solution

beta=0.5
#p=1.17beta
#the value of likelihood
choose(500,200)*(1.17*beta)^200*(1-1.17*beta)^300
#try also
dbinom(200,500,1.17*beta)
#the value of log-likelihood
log(dbinom(200,500,1.17*beta))

#seqencing
beta=seq(0.05,0.85,by=0.001);beta

#define the likelihood function
lh.function=function(beta){dbinom(200,500,1.17*beta)}

#calculate the likelihood function at each beta in the sequence
lh.values=lh.function(beta);lh.values

plot(log(lh.values)~beta,type="l")

#find MLE
max(lh.values)
index=which(lh.values==max(lh.values))
beta.hat=beta[index];beta.hat

#log-likelihood statisitic
LR=2*(log(lh.function(beta.hat))-log(lh.function(0.5)));LR

#p-value
pchisq(LR,1,lower.tail=FALSE)

#critical value
C=qchisq(0.95,1);C

#find 95% C.I.
stat=2*(log(lh.function(beta.hat))-log(lh.function(beta))
w=which(stat<C)
betal=beta[min(w)];betal
beta2=beta[max(w)];beta2