Due: 7th, April 2022

Problems due: 1,2

- 1. A coin is flipped 100 times and 55 heads occurred. Assume that the coin has probability of heads being p.
 - (a) Write an R-function that will compute the value of the likelihood function for any value of p.
 - (b) Plot the likelihood function for $p \in (0, 1)$.
 - (c) Use the optimize function in R to find the maximum likelihood estimate for p.
 - (d) Compute the M.L.E. for p using calculus and see how close is to the answer in the previous step.
 - (e) Do the above steps if the number of observed heads was 30, and 70.
- 2. Suppose we have n samples $X_1, X_2, X_3, \ldots X_n$, from Binomial(N, p). We are told the value of $M = \max\{X_1, X_2, X_3, \ldots X_n\}$.
 - (a) Find the Probability mass function of M.
 - (b) Write an R-function that will compute the value of the likelihood function $L(p \mid M, N, n)$ for any value of p.
 - (c) Suppose M = 30, N = 50, n = 10.
 - i. Plot the likelihood function for $p \in (0, 1)$.
 - ii. Use the optimize function in R to find the maximum likelihood estimate for p.
 - iii. Can you compute the M.L.E. for p using calculus ?
 - (d) Do the previous step if M is now 20, or 40.
- 3. Suppose that a particular gene occurs as one of two alleles (A and a), where allele A has frequency θ in the population. That is, a random copy of the gene is A with probability θ and a with probability 1θ . Since a diploid genotype consists of two genes, the probability of each genotype is given by:

| Genotype | AA | Aa | aa |
|-------------|------------|---------------------|----------------|
| Probability | θ^2 | $2\theta(1-\theta)$ | $(1-\theta)^2$ |

Suppose we test a random sample of the population and find that k] are AA, l are Aa, and m are aa.

- (a) Write an R-function that will compute the value of the likelihood function for any value of θ given k, l, m.
- (b) Suppose k = 10, l = 30, m = 20,
 - i. Plot the likelihood function for $\theta \in (0, 1)$.
 - ii. Use the optimize function in R to find the maximum likelihood estimate for θ .
 - iii. Compute the M.L.E. for θ using calculus and see how close is to the answer in the previous step.
- (c) Do the previous step if the k = l = m = 30.
- 4. Write an R-code that uses the optim or optimize function and solves Bosokovitch's formulation of finding the best line? That is for data points $\{x_i, y_i\}_{i=1}^n$ find m, c that minimizes

$$\sum_{k=1}^{n} |y_k - mx_k - c|.$$
 (1)

5. Read Section 23.4 in the book R For Data Science