

Due: 7th, April 2022

Problems due: 1,2

- A coin is flipped 100 times and 55 heads occurred. Assume that the coin has probability of heads being p .
 - Write an **R**-function that will compute the value of the likelihood function for any value of p .
 - Plot the likelihood function for $p \in (0, 1)$.
 - Use the **optimize** function in **R** to find the maximum likelihood estimate for p .
 - Compute the M.L.E. for p using calculus and see how close is to the answer in the previous step.
 - Do the above steps if the number of observed heads was 30, and 70.
- Suppose we have n samples $X_1, X_2, X_3, \dots, X_n$, from $\text{Binomial}(N, p)$. We are told the value of $M = \{X_1, X_2, X_3, \dots, X_n\}$.
 - Find the Probability mass function of M .
 - Write an **R**-function that will compute the value of the likelihood function $L(p | M, N, n)$ for any value of p .
 - Suppose $M = 30, N = 50, n = 10$.
 - Plot the likelihood function for $p \in (0, 1)$.
 - Use the **optimize** function in **R** to find the maximum likelihood estimate for p .
 - Can you compute the M.L.E. for p using calculus ?
 - Do the previous step if M is now 20, or 40.
- 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 $1 - \theta$. 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 .
 - Write an **R**-function that will compute the value of the likelihood function for any value of θ given k, l, m .
 - Suppose $k = 10, l = 30, m = 20$,
 - Plot the likelihood function for $\theta \in (0, 1)$.
 - Use the **optimize** function in **R** to find the maximum likelihood estimate for θ .
 - Compute the M.L.E. for θ using calculus and see how close is to the answer in the previous step.
 - Do the previous step if the $k = l = m = 30$.
- 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

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

- Read Section 23.4 in the book R For Data Science