BRAC University

CSE427: Machine Learning Summer 2023

Assignment 1: Maximum Likelihood Estimation

July 7, 2023

1st to	2nd to	3rd to	4th to	5th to	6th to	7th toss
T	T	H	T	T	H	T
H	T	T	H	T	T	T
H	H	T	H	H	H	T
H	T	T	H	T	T	T
H	H	T	T	H	T	H
H	H	H	H	T	H	H
H	T	H	T	H	H	T
H	H	H	H	T	T	T
T	T	T	T	T	H	H
H	T	T	H	H	T	T

Machine Learning is the science of learning from experience. Suppose Alice is repeatedly doing an experiment. In each experiment she tosses n coins. She does this experiment m times. In the first round, x_1 coins yielded a head and y_1 coins yielded a tail. Notice that, $x_1 + y_1 = n$. In the second round, x_2 coins yielded a head and y_2 coins yielded a tail. Once again, $x_2 + y_2 = n$. She does this experiment m times. Your job is to estimate the probability p of a coin yielding a head.

1. What is your guess on the value of p?

2. In Maximum Likelihood Estimation, we want to find a parameter p which maximizes all the observations in the dataset. If the dataset is a matrix A, where each row $a_1, a_2, \cdots, a_m$ are individual observations, we want to maximize $P(A) = P(a_1)P(a_2)\cdots P(a_m)$ because individual experiments are independent. Maximizing this is equivalent to maximizing $\log P(A) = \log P(a_1) + \log P(a_2) + \cdots + \log P(a_m)$. Maximizing this quantity is equivalent to minimizing the $-\log P(A) = -\log P(a_1) - \log P(a_2) - \cdots - \log P(a_m)$.

3. Here you need to find out $P(a_i)$ for yourself.

4. If you can do that properly, you will find an equation of the form:

$$-\frac{\log P(A)}{mn} = -\frac{\sum_{i=1}^{m} x_i}{mn} \log p - \frac{\sum_{i=1}^{m} y_i}{mn} \log(1-p)$$

   Now, define $q = \frac{\sum_{i=1}^{m} x_i}{mn}$. Then the equation becomes:

$$-\frac{\log P(A)}{mn} = -q \log p - (1-q)\log(1-p)$$

   Use Pinsker's Inequality or Calculus to show that, $p = q$.

5. What is the value of p for the above dataset given in the table?

6. If you toss 20 coins now, how many coins are most likely to yield a head?