CS561 HW3

Punnawish Thuwajit

September 21, 2023

Student ID : 9084173567

Contents

1. Problem 1: Laws of Probability.

Prove or disprove if the following statements are true in general:

(a) $\mathbb{P}(A^c, B^c) = \mathbb{P}(A^c) - \mathbb{P}(B)$: The statement is false because $\mathbb{P}(A^c, B^c) = \mathbb{P}(A^c) - \mathbb{P}(A, B)$ is true. By the partition law, since, B and B^c are disjoint complements, we obtain

$$\begin{aligned} \mathbb{P}(A^c) &= \mathbb{P}(A^c, B^c) + \mathbb{P}(A^c, B) \\ \mathbb{P}(A^c, B^c) &= \mathbb{P}(A^c) - \mathbb{P}(A^c, B) \end{aligned}$$

And since $\mathbb{P}(A^c, B)$ is not necessarily equal to $\mathbb{P}(B)$, the statement is not true.

(b) $\mathbb{P}(A^c, B^c) \geq \mathbb{P}(A^c) - \mathbb{P}(B)$: The statement is true. We've shown above that $\mathbb{P}(A^c, B^c) = \mathbb{P}(A^c) - \mathbb{P}(A^c, B)$. Moreover, since $A^c \cap B \subseteq B$, we obtain that $\mathbb{P}(A^c, B) \leq \mathbb{P}(B)$. Therefore, $\mathbb{P}(A^c, B^c) \geq \mathbb{P}(A^c) - \mathbb{P}(B)$ as desired.

(c) $\mathbb{P}(A) - \mathbb{P}(B) \leq \mathbb{P}(B^c) - \mathbb{P}(A^c)$: The statement is true. We can rearrange the statement to obtain $\mathbb{P}(A) + \mathbb{P}(A^c) \leq \mathbb{P}(B) + \mathbb{P}(B^c)$, which is equivalent to saying $1 \leq 1$ (because A and A^c are complements, their probability adds up to 1), which is true.

2. Problem 2: The Monty Hall Paradox (Adapted from Vos Savant, 1990).

You're on a game show, and given the choice of three doors, denoted A, B, and C. Behind one is gold, and behind the other two is nothing. Imagine you pick a door. Monty Hall, the host, knows what is behind each door, and opens one of the remaining two doors and shows you there is nothing behind it. He then says to you - "Do you want to pick another door?"

(a) Before you make your decision to move, what is the sample space for this random experiment? WLOG (by symmetry) the first door (A) has the gold, the possible cases are

- You pick door A. Monty Hall can either pick door B or C.
- You pick door B. Monty Hall has to pick door C.
- You pick door C. Monty Hall has to pick door B.

Denote Ω as the sample space where each element is a 2-tuple: the door you choose and the door Monty opens. The sample space can be expressed as follows

$$\Omega = \{(A, B), (A, C), (B, C), (C, B)\}$$

(b) Assume that the gold was placed at random and that your initial choice of door is independent of the gold placement, and that Monty's choice is random among possible empty doors. Specify the probability measure for the outcomes, and use it to compute the probability of winning if the contestant stays or moves. Once again, WLOG (by symmetry), we assume the first door (A) has the gold. Given your initial choice is made uniformly, we can see

$$\mathbb{P}(\{A, B)), (A, C)\}) = \mathbb{P}(\{(B, C)\}) = \mathbb{P}(\{(C, B)\}) = \frac{1}{3}$$

Moreover, Monty picks a door at random, meaning $\mathbb{P}(\{(A, B)\}) = \mathbb{P}(\{(A, C)\}) = \frac{1/3}{2} = \frac{1}{6}$ The events, from Ω, in which you win from keeping your choice is (A, B) and (A, C). As such, the probability you win from keeping your choice is $\mathbb{P}(\{(A, B)\}) + \mathbb{P}(\{(A, C)\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. Its complement, the probability you win from moving is $1 - \frac{1}{3} = \frac{2}{3}$

3. Problem 3 Tertiary (3-class) Classification problem and crossover probabilities

You build a classifier that aims to classify pictures of cats, dogs, and fish. The classifier succeeds at classifying cats as cats with probability $1-\varepsilon$ but incorrectly labels cat pictures as dogs with probability ε, and so on, as specified in the transition diagram below

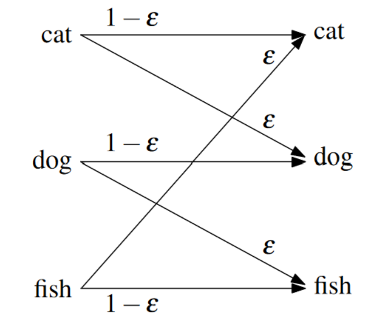

Figure 1: Transition Diagram for Our Classifier

Pictures of cats occur with probability 1/2, pictures of dogs with probability 1/4, and pictures of fish with probability 1/4.

(a) Find, as a function of ε, the probabilities of the labels. We denote cats, dogs, and fish as C, D, F for simplicity

- For cats:

$$\begin{aligned}\mathbb{P}(Y=C) &= \mathbb{P}(Y=C, X=C)+\mathbb{P}(Y=C, X=D)+\mathbb{P}(Y=C, X=F)\\ &= \mathbb{P}(Y=C|X=C)\mathbb{P}(X=C)+\mathbb{P}(Y=C|X=D)\mathbb{P}(X=D)+\mathbb{P}(Y=C|X=F)\mathbb{P}(X=F)\\ &= (1-\varepsilon)(1/2)+(0)(1/4)+(\varepsilon)(1/4)\\ &= 1/2-\varepsilon/4\end{aligned}$$

- For dogs:

$$\begin{aligned}\mathbb{P}(Y=D) &= \mathbb{P}(Y=D, X=C)+\mathbb{P}(Y=D, X=D)+\mathbb{P}(Y=D, X=F)\\ &= \mathbb{P}(Y=D|X=C)\mathbb{P}(X=C)+\mathbb{P}(Y=D|X=D)\mathbb{P}(X=D)+\mathbb{P}(Y=D|X=F)\mathbb{P}(X=F)\\ &= (\varepsilon)(1/2)+(1-\varepsilon)(1/4)+(0)(1/4)\\ &= 1/4+\varepsilon/4\end{aligned}$$

- For fish:

$$\begin{aligned}\mathbb{P}(Y=F) &= \mathbb{P}(Y=F, X=C)+\mathbb{P}(Y=F, X=D)+\mathbb{P}(Y=F, X=F)\\ &= \mathbb{P}(Y=F|X=C)\mathbb{P}(X=C)+\mathbb{P}(Y=F|X=D)\mathbb{P}(X=D)+\mathbb{P}(Y=F|X=F)\mathbb{P}(X=F)\\ &= (0)(1/2)+(\varepsilon)(1/4)+(1-\varepsilon)(1/4)\\ &= 1/4\end{aligned}$$

(b) Suppose the classifier predicts dog. What is the probability that the input was a cat? Fish? Dog? Express your answer as a function of ε. We can use Bayes' rule

$$\mathbb{P}(X=C|Y=D)=\frac{\mathbb{P}(Y=D, X=C)\mathbb{P}(X=C)}{\mathbb{P}(Y=D)}=\frac{(\varepsilon)(1/2)}{1/4+\varepsilon/4}=\frac{2\varepsilon}{1+\varepsilon}$$
$$\mathbb{P}(X=D|Y=D)=\frac{\mathbb{P}(Y=D, X=D)\mathbb{P}(X=D)}{\mathbb{P}(Y=D)}=\frac{(1-\varepsilon)(1/4)}{1/4+\varepsilon/4}=\frac{1-\varepsilon}{1+\varepsilon}$$
$$\mathbb{P}(X=F|Y=D)=\frac{\mathbb{P}(Y=D, X=F)\mathbb{P}(X=F)}{\mathbb{P}(Y=D)}=\frac{(0)(1/4)}{1/4+\varepsilon/4}=0$$

4. Problem 4 Multinomial Probabilities by Enumeration

The multinomial distribution is an extension of the binomial distribution. For n independent trials of an experiment that leads to one of k possible categories (i.e, the roll of a $k = 6$ sided die, n times), the multinomial distribution models the probability of the counts of each category. Imagine that you toss a fair, 6-sided die 7 times.

(a) **Categorical outcomes:** What is Ω? We can enumerate Ω by

```
Omega = [[]] # initiate empty list
for i in range(7): # 7 iterations
    Curr = [] # tack current Omega
    for item in Omega:
        for x in range(1,7): # coin toss from 1 to 6
            Curr.append(item+[x])
    Omega = Curr #update Omega
```

(b) How many total outcomes are there? What is the probability of any given outcome? There are 6^7 (6 outcomes per iteration, 7 iterations) outcomes, each having an equal probability of $\frac{1}{6^7}$ since the dice toss is fair.

(c) Multinomial outcomes. Suppose you are only interested in the number of times each of the categories (sides of the die, 1-6) appeared in the 7 throws. This count or frequency of each category is a multinomial random variable. What is Ω? We can define a 6-tuple $(a_1, a_2, \dots, a_6)$ as the number of times each number is rolled. Since there are 7 dice, $a_1 + a_2 + \cdots + a_6 = 7$. Moreover, a_i must take a non-negative integral value. Therefore,

$$\Omega = \left\{ (a_1, a_2, \dots, a_6) : a_i \in \mathbb{Z}_0^+, a_1 + a_2 + \cdots + a_6 = 7 \right\}$$

(d) Write a script to convert the outcomes from (a) into this format, keeping track of how many outcomes in (a) map to a single multinomial outcome.

```
def convert(dice_roll):
    count = {i:0 for i in range(1,6)}
    for i in dice_roll:
        count[i]+=1
    return tuple([count[i] for i in range(1,6)])
New_Omega = dict()
for dice_roll in Omega:
    count = convert(dice_roll)
    if count in New_Omega.keys():
        New_Omega[count]+=1
    else:
        New_Omega[count]=1
```

Where we take in **Omega** from (a) and computes **New_Omega**, a dictionary in the form of $(a_1, a_2, \dots, a_6)$: its count, for each possible nonnegative integral tuple $(a_1, a_2, \dots, a_6)$ that satisfies $a_1 + a_2 + \cdots + a_6 = 7$.

(e) How many possible multinomial outcomes are there? We can analytically compute the integral solution of $a_1+a_2+\cdots+a_6 = 7$ with the stars and bars method, yielding $\binom{7+6-1}{6-1} = \binom{12}{5} = 792$ solutions. Enumerating the numbers of keys in **New_Omega** yields the same answer.

(f) What are the most probable multinomial outcomes, and what is their probability? $(2, 1, 1, 1, 1, 1)$ (and its permutations) is the most probable, each having a probability of $\frac{2520}{6^7}$.

(g) Derive a general expression for the total number of multinomial outcomes given k categories and n trials. Similar to (c), we can treat each possible number of outcomes as an integral solution to

$$a_1 + a_2 + \cdots + a_k = n$$

where each a_i is a nonnegative integer. We can apply the stars and bars approach, that is out of $n+k-1$ stars, we pick $k-1$ of them to replace with bars. The stars from start to the first is a_1, from first to second is a_1, and so on, each selection corresponding to a unique $(a_1, a_2, \ldots, a_k)$ tuple. As such, we can compute the number of solutions to be $\binom{n+k-1}{k-1}$