Probability & Statistics for EECS: Homework #02

Due on Oct 22, 2023 at 23:59

Name: **Runkang Yang**
Student ID: 2022533080

Problem 1

1. (a)

   Suppose that we use A to denote that Alice actually sent a 1, and B to denote that Bob receives a 1. Then we have

$$\begin{aligned} P(A|B) &= \frac{P(A)P(B|A)}{P(B)} \\ &= \frac{P(A)P(B|A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.5 * 0.9}{0.9 * 0.5 + 0.05 * 0.9} \\ &= \frac{18}{19} \end{aligned}$$

   So given that Bob receives a 1, the probability that Alice actually sent a 1 is $\frac{18}{19}$.

   (b)

   Similarly, we use A to denote that Alice actually sent 111, and B to denote that Bob receives 110. Then we have

$$\begin{aligned} P(A|B) &= \frac{P(A)P(B|A)}{P(B)} \\ &= \frac{P(A)P(B|A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.5 * 0.9 * 0.9 * 0.1}{0.9 * 0.9 * 0.1 * 0.5 + 0.05 * 0.05 * 0.95 * 0.5} \\ &= \frac{648}{667} \end{aligned}$$

   So given that Bob receives 110, the probability that Alice intended to convey a 1 is $\frac{648}{667}$.

Problem 2

1. (a)

Suppose that we use A to denote that Fred has the disease, and B to denote that he tests positive on all n of the n tests. Then we have

$$\begin{aligned}
P(B|A) &= a^n \\
P(B|A^c) &= b^n \\
P(A|B) &= \frac{P(A)P(B|A)}{P(B)} \\
&= \frac{P(A)P(B|A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\
&= \frac{P(D)a^n}{P(D)a^n + [1 - P(D)]b^n} \\
&= \frac{pa^n}{pa^n + qb^n}
\end{aligned}$$

So given that Fred tests positive on all n of the n tests, the posterior probability that Fred has the disease is $\frac{pa^j}{pa^j+qb^j}$.

(b) Similarly, we use A to denote that Fred has the disease, and B to denote that he tests positive on all n of the tests. Then we have

$$\begin{aligned}
P(B|A) &= P(B|A,G)P(G|A) + P(B|A,G^c)P(G^c|A) \\
&= P(B|A,G)P(G) + P(B|A,G^c)P(G^c) \\
&= \frac{1}{2} + \frac{1}{2}a_0^n \\
P(B|A^c) &= P(B|A^c,G)P(G|A^c) + P(B|A^c,G^c)P(G^c|A^c) \\
&= P(B|A^c,G)P(G) + P(B|A^c,G^c)P(G^c) \\
&= \frac{1}{2} + \frac{1}{2}b_0^n \\
P(A|B) &= \frac{P(A)P(B|A)}{P(B)} \\
&= \frac{P(A)P(B|A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\
&= \frac{P(D)(\frac{1}{2} + \frac{1}{2}a_0^n)}{P(D)(\frac{1}{2} + \frac{1}{2}a_0^n) + [1 - P(D)](\frac{1}{2} + \frac{1}{2}b_0^n)} \\
&= \frac{p(\frac{1}{2} + \frac{1}{2}a_0^n)}{p(\frac{1}{2} + \frac{1}{2}a_0^n) + q(\frac{1}{2} + \frac{1}{2}b_0^n)}
\end{aligned}$$

Problem 3

1. Suppose that we use A to denote that a new email has just arrived, and it include the 23rd, 64th, and 65th words or phrases on the list So A is W_1^c Then we have

$$
\begin{aligned}
P(A|spam) &= (1-p_1)(1-p_2)...p_{23}...p_{64}...p_{65}...(1-p_{100}) \\
P(A|not\ spam) &= (1-r_1)(1-r_2)...r_{23}...r_{64}...r_{65}...(1-r_{100}) \\
P(spam|A) &= \frac{P(spam)P(A|spam)}{P(A)} \\
&= \frac{P(A|spam)P(spam)}{P(A|spam)P(spam)+P(A|not\ spam)P(not\ spam)} \\
&= \frac{p(1-p_1)(1-p_2)...p_{23}...p_{64}...p_{65}...(1-p_{100})}{(1-p_1)(1-p_2)...p_{23}...p_{64}...p_{65}...(1-p_{100})+(1-p)(1-r_1)(1-r_2)...r_{23}...r_{64}...r_{65}...(1-r_{100})}
\end{aligned}
$$

So the condition probability that the new email is spam is

$$
\frac{p(1-p_1)(1-p_2)...p_{23}...p_{64}...p_{65}...(1-p_{100})}{(1-p_1)(1-p_2)...p_{23}...p_{64}...p_{65}...(1-p_{100})+(1-p)(1-r_1)(1-r_2)...r_{23}...r_{64}...r_{65}...(1-r_{100})}
$$

.

Problem 4

1. Suppose that we first choose the door 1. Define C_i as a car is behind door i, i $\in$ {1,2,3}, so we can conclude that $P(C_1) = p_1, P(C_2) = p_2, P(C_3) = p_3$. And we use A to denote that I winning a car, we can get the below formula by LOTP :

   $P(A) = P(A|C_1)P(C_1) + P(A|C_2)P(C_2) + P(A|C_3)P(C_3)$

   Now let's consider the situation that I choose to switch my choice. So if the car is behind the other doors, the probability of winning a car is 1, while if the car is behind the door 1, the probability of winning a car is 0. So $P(A) = 0 \times 1 + 1 \times p_2 + 1 \times p_3 = p_2 + p_3$.

   And if I don't choose to switch my choice, if the car is behind the other doors, the probability of winning a car is 0, while if the car is behind the door 1, the probability of winning a car is 1. So $P(A) = 1 \times p_1 + 0 \times p_2 + 0 \times p_3 = p_1$.

   So let's compare p_1 and $p_2 + p_3$, since $p_2 + p_3 = 1 - p_1$, so when $p_1\ in(0, \frac{1}{2})$, then $p_2 + p_3 = 1 - p_1 \in (\frac{1}{2}, 1) > p_1$, so we're expected to switch the initial choice, when $p_1 = \frac{1}{2}$, each strategy is both acceptable, when $p_1\ in(\frac{1}{2}, 1)$, then $p_2 + p_3 = 1 - p_1 \in (0, \frac{1}{2}) < p_1$, so we're expected not to switch the initial choice.

   And the above only discuss the situation that we first choose the door 1, it seems like we should also consider other two situations in detail, which is first choose the door 2 or first choose the door 3. However, if we take a closer look at this issue, we can combine the two conditions mentioned in the question $\begin{cases} p_1 + p_2 + p_3 = 1 \\ p1 \geq p_2 \geq p_3 > 0 \end{cases}$ to conclude that $p_2, p_3 \in (0, \frac{1}{3}) < \frac{1}{2}$.

   So similarly, if we first choose the door 2, and if I switch my choice, we get $P(A) = 0 \times p_2 + 1 \times p_1 + 1 \times p_3 = p_1 + p_3$, if I don't choose to switch my choice, we get $P(A) = 1 \times p_2 + 0 \times p_1 + 0 \times p_3 = p_2$, since $p_2 < \frac{1}{2}$, so $p_1 + p_3 > p_2$ and switch the initial choice is the best strategy. If we first choose the door 3, and if I switch my choice, we get $P(A) = 0 \times p_3 + 1 \times p_1 + 1 \times p_2 = p_1 + p_2$, if I don't choose to switch my choice, we get $P(A) = 1 \times p_3 + 0 \times p_1 + 0 \times p_2 = p_3$, since $p_3 < \frac{1}{2}$, so switch the initial choice is the best strategy, too.

   To draw a conclusion ,the answer to the question is: If I first choose the door 1, and if $p_1 \in (0, \frac{1}{2})$, we're expected to switch the initial choice, if $p_1 \in (\frac{1}{2}, 1)$, we're expected not to switch the initial choice, if $p_1 = \frac{1}{2}$, two strategies are both acceptable.
   If I first choose the door 2 or 3, we're expected to switch the initial choice.

Problem 5

1. (a)

   Define C_i as a car is behind door i, i $\in$ {1,2,3}, so we can conclude that $P(C_1) = P(C_2) = P(C_3) = \frac{1}{3}$. And we use A to denote that winning a car by switching, we can get the below formula by LOTP :

$$\begin{aligned} P(A) &= P(A|C_1)P(C_1) + P(A|C_2)P(C_2) + P(A|C_3)P(C_3) \\ &= 0 \times \frac{1}{3} + 1 \times \frac{1}{3} + 1 \times \frac{1}{3} \\ &= \frac{2}{3} \end{aligned}$$

   (b)

   Define D_i as Monty open door i, i $\in$ {1,2,3}, and given that Monty opens door 2, if the strategy of always switching succeeds, it means that we first choose door 1 and then switch to door 3, or first choose door 3 and then switch to door 1 without loss of generality. so we can conclude that

$$\begin{aligned} P(D_2|C_3) &= 1 \\ P(D_2) &= P(D_2|C_1)P(C_1) + P(D_2|C_2)P(C_2) + P(D_2|C_3)P(C_3) \\ &= p \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} \\ &= \frac{p+1}{3} \\ P(C_3|D_2) &= \frac{P(D_2|C_3)P(C_3)}{P(D_2)} \\ &= \frac{P(D_2|C_3)P(C_3)}{P(D_2|C_1)P(C_1) + P(D_2|C_2)P(C_2) + P(D_2|C_3)P(C_3)} \\ &= \frac{1 \times \frac{1}{3}}{\frac{p+1}{3}} \\ &= \frac{1}{p+1} \end{aligned}$$

   So given that Monty opens door 2, the probability that the strategy of always switching succeeds is $\frac{1}{p+1}$.

   (c)

   Based on the conditions given in the question, Monty opens door 2 with probability p,so he opens door 3 with probability 1-p. And through our analysis, the situations in (c) and (b) questions can be seen as relevant. So similarly, given that Monty opens door 3, if the strategy of always switching succeeds, it means that we first choose door 1 and then switch to door 2, or first choose door 2 and then switch to door 1 without loss of generality. We just use 1-p to substitude p,and conclude that

$$\begin{aligned} P(D_3|C_2) &= 1 \\ P(D_3) &= \frac{2-p}{3} \\ P(C_2|D_3) &= \frac{P(D_2|C_3)P(C_2)}{P(D_3)} \\ &= \frac{1}{2-p} \end{aligned}$$

   So given that Monty opens door 3, the probability that the strategy of always switching succeeds is $\frac{1}{2-p}$.

Problem 6

1. Through our analysis, if the nth trial is successful, it means Treatment A is assigned for the trial and success, whose probability is $a_a n$, or it means that Treatment B is assigned for the trial and success, whose probability is $(1-a_n)b$, and according to the addition principle, we add the two result up, and find that

$$\begin{aligned} p_n &= aa_n + (1-a_n)b \\ &= (a-b)a_n + b \end{aligned}$$

If Treatment A is assigned on the $(n+1)$th trial, it means that nth trial is also successful with Treatment A and Treatment A is assigned on the $(n+1)$th trial, whose probability is aa_n, or it means that the nth trial is failure with Treatment B and Treatment A is assigned on the $(n+1)$th trial, whose probability is $(1-a_n)(1-b)$, and according to the addition principle, we add the two result up, and find that

$$\begin{aligned} a_{n+1} &= aa_n + (1-a_n)(1-b) \\ &= aa_n + 1 - b - a_n + a_n b \\ &= (a+b-1)a_n + 1 - b \end{aligned}$$

(b)

Let n=n+1, and use the result from (a), we can find that

$$\begin{aligned} p_{n+1} &= (a-b)a_{n+1} + b \\ &= (a-b)[(a+b-1)a_n + 1 - b] + b \\ &= (a-b)[(a+b-1)\frac{p_n - b}{(a-b)} + 1 - b] + b \\ &= (a+b-1)p_n - b(a+b-1) + (a-b)(1-b) + b \\ &= (a+b-1)p_n - ab - b^2 + b + a - ab - b + b^2 + b \\ &= (a+b-1)p_n + a + b - 2ab \end{aligned}$$

(c) Since $p_{n+1} = (a+b-1)p_n + a + b - 2ab$. We can assume that $A = (a+b-1)$, and $B = a+b-2ab$, so we get $p_{n+1} = Ap_n + B$, and there exist C to satisfy $p_{n+1} + C = A(p_n + C)$,which is $p_{n+1} = Ap_n + C(A-1)$. So $C(A-1) = B$, then $C = \frac{B}{A-1}$.
And we can find that $p_n + C$ forms a geometric sequence since $\frac{p_{n+1}+C}{p_n+C} = A$ which we may denote that $p_n + C = (p_1 + C)A^{n-1}$. So,

$$\begin{aligned} p_n &= (p_1 + C)A^{n-1} - C \\ &= [(a-b)a_1 + b + \frac{a+b-2ab}{a+b-2}](a+b-1)^{n-1} - \frac{a+b-2ab}{a+b-2} \\ &= (\frac{a-b}{2} + b + \frac{a+b-2ab}{a+b-2})(a+b-1)^{n-1} - \frac{a+b-2ab}{a+b-2} \\ &= [\frac{(a+b)(a+b-2) + 2(a+b-2ab)}{2(a+b-2)}](a+b-1)^{n-1} - \frac{a+b-2ab}{a+b-2} \\ &= \frac{(a-b)^2(a+b-1)^{n-1}}{2(a+b-2)} - \frac{a+b-2ab}{a+b-2} \end{aligned}$$

Since $a+b<2$, then $a+b-1<1$, and $\lim_{n\to\infty}(a+b-1)^{n-1} = 0$, so $\lim_{n\to\infty} p_n = -\frac{a+b-2ab}{a+b-2}$.