241880334 闵振昌 25FA2

Problem 1

The Prosecutor's Fallacy

For conditional probabilities we have:

$$Pr(G|T) = \frac{Pr(G \cap T)}{Pr(T)} = \frac{Pr(T|G)Pr(G)}{Pr(T)}$$

So if $Pr(G|T) = Pr(T|G)$, then $\frac{Pr(G)}{Pr(T)} = 1$ So if $Pr(G|T) = Pr(T|G)$, then $Pr(G) = Pr(T)$

And as this is an equation, given $Pr(G) = Pr(T)$ we can infer that $Pr(G|T) = Pr(T|G)$

Sure Thing Principle

I agree, as $C \cup C^c = \Omega$ is a certain event, covering all the possible situations. And given that x outperforms y in all situations, then preferring x to y can be inferred.

The Prisoner's Dilemma

The problem is that after he knew who would be released other than him, the probability turns into a conditional probability instead of a normal one. Without asking, the probability of himself being released is $\frac{2}{3}$ and that is correct.

Let the asking prisoner be A and other two be B and C. The probability of A being released after asking can be calculated as:

$$Pr(A|B \text{ getting out}) = Pr(A|C \text{ getting out}) = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6}} = \frac{2}{3}$$

The Monty Hall Problem

1. (1) There're 3 situations in general, if the car is in the chosen door, then the probability of A getting the car in the third one is 0. But if in other 2 situations when the chosen door is goat, we would definitely get one car in the third door, so it is easy to get $p = \frac{2}{3}$

(2) We can know if we see Bill, we either choose Nan's door or we chose a car's door and the host reveals Bill. If we've chose Nan's door, then the probability of getting the car is 1 If we've chose a car's door, then the probability of getting the car is 0. Probability of seeing Bill and car in third door can be inferred as only when we've chosen Nan's door initially, and we'll definitely get the car in the third door cuz Bill is already revealed. So the conditional probability is
$\frac{Pr(\text{We see Bill and car in third door})}{Pr(\text{We see Bill})} = \frac{\frac{1}{3}}{\frac{b}{3} + \frac{1}{3}} = \frac{1}{b+1}$

(3)Firstly we need to make sure the one we chose initially contains a goat, not a car, or we will never get a car in the third door, and this probability is $\frac{2}{3}$ And then the probability of getting the car in the third door is $\frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$ And the probability of the host revealing a goat is $\frac{1}{3} + \frac{2}{3} \times \frac{1}{2} = \frac{2}{3}$ So $p = \frac{\frac{1}{3}}{\frac{2}{3}} = \frac{1}{2}$

2. In subproblem 1.(2), we can know that the conditional probability of getting the car in the third door is $\frac{1}{b+1}$ And assume we get Bill for $b \geq \frac{1}{2}$, we can easily know now $p \in [\frac{1}{2}, \frac{2}{3}]$, so by adjusting probability b we can set a protocol for every p. And as $p \geq \frac{1}{2} > \frac{1}{3}$, it's always better to change the chosen door to the third one as it has bigger probability.

3. p=0: When the first chosen door is car, the presenter opens a second door. When it is a goat, the presenter rewards you of the goat in your first chosen door.

p=1: When the first chosen door is goat, the presenter opens a second door concealing another goat. When it is a car, the presenter rewards you of the car in your first chosen door.

for any $\alpha \in [0, 1]$, we set the protocol as: (1) If the first chosen door is car, then the presenter rewards you of the car in your first chosen door. (2) If the first chosen door is goat, then the presenter opens the other goat for probability α. In this way the conditional probability $p = \alpha$

Problem 2

Gambler's Ruin

Set event $P(k)$ as the gambler getting to \$n with initial wealth \$k. By analysis we know:

$$P(k) = P(k-1) \times (1-p) + P(k+1) \times p$$

With calculus knowledge we solve this equation to get the formula:

$$p \times r^2 - r + (1-p) = 0$$

$$r = \frac{1 \pm |2p - 1|}{2p}$$

If $p = \frac{1}{2}$, then $r = 1$, the solution of equation:

$$P(k) = A + B \times k$$

Using $P(0) = 0$ and $P(n) = 1$ we know $P(k) = \frac{k}{n}$

If $p \neq \frac{1}{2}$, then $r_1 = 1$, $r_2 = \frac{1-p}{p}$ Using $P(0) = 0$ and \$P(n)=1\$to get solution as:

$$P(k) = \frac{(\frac{p}{1-p})^k - (\frac{p}{1-p})^n}{1 - (\frac{p}{1-p})^n}$$

In conclusion, the probability of the gambler getting to \$n with initial wealth \$k is:

$$P(k) = \begin{cases} \frac{k}{n} & , \; p = \frac{1}{2} \\ \frac{(\frac{p}{1-p})^k - (\frac{p}{1-p})^n}{1 - (\frac{p}{1-p})^n} & , \; p \neq \frac{1}{2} \end{cases}$$

Problem 3

Certainty and Independency

1. Easily we know if A is independent of itself, then $Pr(A \cap A) = Pr(A)$. By independence we know $Pr(A \cap A) = Pr(A) \times Pr(A) = Pr(A)^2 = Pr(A)$ Solving this we know $Pr(A) = 0$ or $Pr(A) = 1$

2. If $Pr(A) = 0$, for any event B, $Pr(B \cap A) = Pr(B) \times Pr(A) = Pr(B) \times 0 = 0$ If $Pr(A) = 1$, for any event B, $Pr(B \cap A) = Pr(B) \times Pr(A) = Pr(B) \times 1 = Pr(B)$ Therefore we know A is independent of any event B if A is a certain event or an impossible event.

Out of Service

Firstly we define random variables X as the number of active voice users, and Y as the number of active data users. Both of them follows a binomial distribution with parameters n_1, p_1 and n_2, p_2.

And the total data rate R is the sum of voice data rate R_V and data data rate R_D.

$$R = R_V + R_D = r_1 \times X + r_2 \times Y$$

We need to get the probability that $P(R > c) = P(r_1 X + r_2 Y > c)$ As X and Y are independent, we know:

$$P(R > c) = \sum_{r_1 x + r_2 y > c} P(X = x)P(Y = y) = \sum_{r_1 x + r_2 y > c} \binom{n_1}{x} p_1^x (1 - p_1)^{n_1 - x} \binom{n_2}{y} p_2^y (1 - p_2)^{n_2 - y}$$

Pairwise Independencies

Let's focus on 1 and 2 first. Now given that $Pr(A|B) = Pr(A)$, then $Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} = Pr(A)$, and we have $Pr(A \cap B) = Pr(A)P(B)$, which is identical to equation 2. In this way we prove that equation 1 and 2 are equivalent.

Then with equation 3 we get $Pr(A \cap B^c) = Pr(A)Pr(B^c) = Pr(A)(1 - Pr(B)) = Pr(A) - Pr(A)Pr(B)$ So $Pr(A)Pr(B) = Pr(A) - Pr(A)Pr(B^c) = Pr(A|B)$, so equation 3 is equivalent to equation 2.

Limited Independence

$Pr(A_1) = \frac{4}{36} = \frac{1}{9}$ $Pr(A_2) = \frac{1}{2}$ $Pr(A_3) = \frac{1}{2}$ $Pr(A_1 \cap A_2 \cap A_3) = \frac{1}{36}$ And coincidentally $Pr(A_1 \cap A_2 \cap A_3) = Pr(A_1) \times Pr(A_2) \times Pr(A_3) = P\frac{1}{36}$ But, are they mutually independent? No, they are not. As $Pr(A_2 \cap A_3) = \frac{1}{6} \neq Pr(A_2) \times Pr(A_3) = \frac{1}{4}$

Galton's Paradox

When flipping 3 coins, there're 8 possible outcomes in total: HHH, HHT, HTH, HTT, THH, THT, THH, TTT. As two of them illustrate that they're all alike, the probability should be $\frac{1}{4}$ instead of $\frac{1}{2}$

The reason that this inferrence goes wrong is that indeed there're definitely at least 2 coins alike, but we can't assume that it is not concluding the third one. If the original two are not alike, no matter what we get on the third, we won't get 3 alike coins. So if we wonna calculate in this way, we must multiply the probability that the original two are alike coins, and that is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

Conditional Independence

To prove conditional independence never implies or being implied by independence, we set $\Omega = \{1, 2, 3, 4, 5, 6\}$ Set $A = \{1, 2, 3\}$, $B = \{1, 5, 6\}$, $C = \{1, 6\}$ Conditional independence can be proved as

$$Pr(A \cap B | C) = Pr(A|C) \times Pr(B|C) = \frac{1}{2} \times 1 = \frac{1}{2}$$

While independence can not be proved:

$$Pr(A) \times Pr(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

While $Pr(A \cap B) = \frac{1}{6} \neq Pr(A) \times Pr(B)$

Another example we set $\Omega = \{1, 2, 3, 4\}$ Set $A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 2, 3\}$ Easily we get A and B are independent, but given the condition C, they're not conditional independent.

If $Pr(C) = 0$ or $Pr(C) = 1$, then $Pr(A \cap B | C) = Pr(A \cap B) = Pr(A) \times Pr(B)$

Problem 4

Independent Sets

Choose a subset S of V uniformly at random, each vertex has probability p being selected, and let X be the number of vertices in S, Y be the number of edges in S.

To form independent set, we delete one vertex of every edge in S, getting a new subset S', and the vertices of S' is no more than $X - Y$.

Then we can infer that:

$$E(S') \geq E(X - Y) = E(X) - E(Y) = np - mp^2$$

To maximize this, we can know that when $p = \frac{n}{2m}$, $E(S')$ reaches its maximum value $\frac{n^2}{4m}$ Now we know the expectation of vertices in independent set $|S'| \geq \frac{n^2}{4m}$