

CS203B

Mathematics For Computer
Science -III (Probability)
Indian Institute of Technology, Kanpur

Assignment 2

Name:
Anjali Rana (190147)

Date of Submission:
9th April 2021

Question 1

[5 points] We toss a fair coin n times. Prove that the length of the longest sequence of consecutive heads will not be more than $2\log n$ with probability at least $1 - \frac{1}{n}$.

Solution

To Prove: Longest sequence of consecutive heads will not be more than $2\log n$ with probability at least $1 - \frac{1}{n}$

Proof. Let $N(k)$ = Number of sequences of length n in which longest sequence of consecutive heads is atleast k .

For calculating this count, we select k places and allow anything to be present at the left $n - k$ places. But in doing so, we calculate many cases more than once.

For example: Suppose $n = 5$ and $k = 2$. If we fix k places from index 2 to 3, then we count this case: THHHT here once. Then when we fix k places from index 3 to 4, we again count this case when we allow anything to be present at the rest positions. So, we saw that there are a lot of overlapping cases here.

Now, ways of choosing consecutive k places from n places are $n - k + 1$ and ways of allowing anything to be present at left $n - k$ places are 2^{n-k} . Hence we have,

$$N(k) \leq (n - k + 1)2^{n-k}$$

Total number of possible sequences are 2^n .

Let A be the event in which longest sequence of consecutive heads is atleast k . So,

$$P(A) = \frac{\text{Favourable Cases}}{\text{Total Cases}}$$

$$\leq \frac{(n - k + 1)2^{n-k}}{2^n}$$

$$\text{Put } k = 2\log n + 1$$

$$\leq \frac{(n - 2\log n)2^{n-2\log n-1}}{2^n}$$

$$\leq \frac{(n - 2\log n)}{2n^2}$$

$$\leq \frac{(n)}{2n^2}$$

$$\leq \frac{1}{2n}$$

$$\leq \frac{1}{n}$$

$$1 - P(A) \leq 1 - \frac{1}{n}$$

Now,

$1 - P(A)$ = Probability that longest sequence of heads is less than or equal to $(k - 1)$ (i.e. $2\log n$).

Hence proved,

$$P(\text{longest sequence of heads is no more than } 2\log n) \leq 1 - \frac{1}{n}$$

□

Question 2

[5 + 3 points] There are two bridges from town A to town B and two bridges from town B to town C . Each of the four bridges is blocked by snow with probability p , independent of the others.

- Find the probability that there is an open bridge from A to B given that there is no open route from A to C .
- Say, in addition, there is a direct bridge from A to C ; this bridge being blocked with probability p independently of the others. Find the required conditional probability mentioned above.

Solution

We define the following events:

$XorY$: There exists an open route from X to Y .

$XnorY$: There is no open route from X to Y .

$XobY$: There exists an open bridge from X to Y .

$XnobY$: There is no open bridge from X to Y .

(a)

$$P(AobB/AnorC) = \frac{P(AobB)P(AnorC/AobB)}{P(AnorC)}$$

(i) As per the question,

$$P(AnobB) = P(\text{both bridges from } A \text{ to } B \text{ are blocked (independently)}) = p^2$$

$$\text{So, } P(AobB) = 1 - P(AnobB) = 1 - p^2 \quad \dots(1)$$

(ii) $P(AnorC/AobB)$:

We have an open bridge from A to B . If there must be no open route from A to C then every bridge from B to C must be closed. Hence,

$$P(AnorC/AobB) = P(\text{both bridges from } B \text{ to } C \text{ are closed}) = p^2 \quad \dots(2)$$

(iii)

$$\begin{aligned}P(AnorC) &= 1 - P(AorC) \\&= 1 - P(AobB)P(BobC) \\&= 1 - (1 - p^2)(1 - p^2) \quad \dots(3)\end{aligned}$$

(Since $P(AobB) = P(BobC) = 1 - p^2$ (from (1))

Put (1), (2) and (3) in (a):

$$\begin{aligned}P(AobB/AnorC) &= \frac{(1 - p^2)(p^2)}{(1 - (1 - p^2)(1 - p^2))} \\&= \frac{1 - p^2}{2 - p^2}\end{aligned}$$

(b)

$$P(AobB/AnorC) = \frac{P(AobB)P(AnorC/AobB)}{P(AnorC)}$$

(i) As per the question,

$P(AnobB) = P(\text{both bridges from } A \text{ to } B \text{ are blocked (independently)}) = p^2$ (No effect of added bridge from A to C)

So, here also $P(AobB) = 1 - P(AnobB) = 1 - p^2 \quad \dots(1)$

(ii) $P(AnorC/AobB)$:

We have an open bridge from A to B . So, if we want that there must be no open route from A to C then the direct bridge from A to C must be closed (independently from others) AND also every bridge from B to C must be closed (since there is an open

bridge from A to B and we don't want an open route from A to C). Hence,

$$\begin{aligned}
 P(A \text{ nor } C / A \text{ ob } B) &= P(\text{both bridges from } B \text{ to } C \text{ are closed AND direct bridge from } A \text{ to } C \text{ is closed}) \\
 &\quad (\text{Since direct route is independent of route via } B, \text{ we have):} \\
 &= P(\text{both bridges from } B \text{ to } C \text{ are closed}) P(\text{direct bridge from } A \text{ to } C \text{ is closed}) \\
 &= p^2 p \\
 &= p^3 \quad \dots (2)
 \end{aligned}$$

(iii)

$$\begin{aligned}
 P(A \text{ nor } C) &= P(\text{direct bridge from } A \text{ to } C \text{ is closed AND route from } A \text{ to } C \text{ via } B \text{ is closed}) \\
 &\quad (\text{Since direct route is independent of route via } B, \text{ we have):} \\
 &= P(\text{direct bridge from } A \text{ to } C) P(\text{route from } A \text{ to } C \text{ via } B \text{ is closed}) \\
 &= p(1 - P(A \text{ or } C)) \\
 &= p(1 - P(A \text{ ob } B) P(B \text{ ob } C)) \\
 &= p(1 - (1 - p^2)(1 - p^2)) \quad \dots (3)
 \end{aligned}$$

Put (1), (2) and (3) in (a):

$$\begin{aligned}
 P(A \text{ ob } B / A \text{ nor } C) &= \frac{(1 - p^2)p^3}{p(1 - (1 - p^2)(1 - p^2))} \\
 &= \frac{1 - p^2}{2 - p^2}
 \end{aligned}$$

Hence, in both the cases, we get the same probability.

Question 3

[7points] In a male-dominated sexist society, there are n married couples. Each married couple aims at having at least one male child. So each couple practises the following rule: Keep on producing children until either the number of children becomes 10 or one male child is born. What will be the ratio of males and females in the next generation? Assume that each child is going to be male or female with equal probability. Give mathematical reason in support of your answer.

Solution

Suppose we define random variable X = number of males in society.

Let X_i = number of males in family i .

So, $X = \sum_{i=1}^n X_i$

Probability distribution function of X_i :

(i) $P(X_i = 0)$ = all 10 females = $(\frac{1}{2})^{10}$

(ii) $P(X_i = 1)$ calculated in (2)

Now we have,

$$E[X] = E\left[\sum_{i=1}^n X_i\right]$$

Using linearity of expectation,

$$E[X] = \sum_{i=1}^n E[X_i]$$

\therefore all X_i are identical,

$$E[X] = nE[X_1]$$

Now,

$$E[X_1] = \sum_{x \in \text{Range}(X_1)} P(X_1 = x)x$$

Since X_1 can take only values 0 or 1,

$$E[X_1] = P(X_1 = 1).1 + P(X_1 = 0).0 \quad \dots(1)$$

Now, let M_i be the event that the i_{th} child is male (i.e. first $(i - 1)$ children are females).

$$\begin{aligned} P(X_1 = 1) &= \sum_{i=1}^{10} M_i \\ &= \sum_{i=1}^{10} \left(\frac{1}{2}\right)^i \\ &\quad \text{(Using summation of GP with common ratio} = \frac{1}{2}\text{)} \\ &= \frac{\frac{1}{2}(1 - (\frac{1}{2})^{10})}{1 - \frac{1}{2}} \\ &= 1 - \left(\frac{1}{2}\right)^{10} \quad \dots(2) \end{aligned}$$

Putting (2) in (1),

$$E[X_1] = 1 - \left(\frac{1}{2}\right)^{10}$$

So,

$$E[X] = nE[X_1] = n(1 - (\frac{1}{2})^{10}) \quad \dots(3)$$

Now, suppose we define random variable Y = number of females in society.

Let Y_i = number of females in family i .

So, $Y = \sum_{i=1}^n Y_i$

Probability distribution function of Y_i :

$$P(Y_i = y < 10) = \text{first } y \text{ females and } (y + 1)_{th} \text{ child as male} = \left(\frac{1}{2}\right)^{y+1} \quad \dots(4)$$

$$P(Y_i = 10) = \text{all 10 children are female} = \left(\frac{1}{2}\right)^{10} \quad \dots(5)$$

Now we have,

$$E[Y] = E\left[\sum_{i=1}^n Y_i\right]$$

Using linearity of expectation,

$$E[Y] = \sum_{i=1}^n E[Y_i]$$

\because all Y_i are identical,

$$E[Y] = nE[Y_1] \quad \dots(6)$$

Now,

$$E[Y_1] = \sum_{y \in \text{Range}(Y_1)} P(Y_1 = y)y$$

Since Y_1 can take values 0 to 10,

$$E[Y_1] = \sum_{y=0}^{10} P(Y_1 = y)y$$

$$E[Y_1] = \sum_{y=0}^{10} P(Y_1 = y)y$$

Now using probability distribution function from (4) and (5) we get,

$$E[Y_1] = \sum_{y=0}^{10} P(Y_1 = y)y$$

$$E[Y_1] = 10\left(\frac{1}{2}\right)^{10} + \sum_{y=0}^9 \left(\frac{1}{2}\right)^{y+1}y \quad \dots(7)$$

Suppose,

$$A = \sum_{y=1}^9 \left(\frac{1}{2}\right)^{y+1}y \quad \dots(i)$$

Multiply both sides with $\frac{1}{2}$

$$\frac{1}{2}A = \sum_{y=1}^9 \left(\frac{1}{2}\right)^{y+2}y \quad \dots(ii)$$

Subtract ii from i

$$\frac{1}{2}A = \sum_{y=1}^9 \left(\frac{1}{2}\right)^{y+1} - \left(\frac{1}{2}\right)^{11} 9$$

$$A = \left(\sum_{y=1}^9 \left(\frac{1}{2}\right)^y\right) - 9\left(\frac{1}{2}\right)^{10}$$

$$A = \left(1 - \left(\frac{1}{2}\right)^9\right) - 9\left(\frac{1}{2}\right)^{10}$$

Putting A in (7),

$$E[Y_1] = 10\left(\frac{1}{2}\right)^{10} + \left(1 - \left(\frac{1}{2}\right)^9\right) - 9\left(\frac{1}{2}\right)^{10}$$

$$E[Y_1] = 1 - \left(\frac{1}{2}\right)^{10}$$

$$E[Y] = nE[Y_1] \quad \dots(\text{from 6})$$

$$E[Y] = n\left(1 - \left(\frac{1}{2}\right)^{10}\right) \quad \dots(8)$$

From (3) and (8)

$$\frac{E[X]}{E[Y]} = \frac{n\left(1 - \left(\frac{1}{2}\right)^{10}\right)}{n\left(1 - \left(\frac{1}{2}\right)^{10}\right)}$$

$$\frac{E[X]}{E[Y]} = 1$$

Question 4

[5 points] X is a continuous random variable with probability distribution function $F_X(x) := P(X \leq x)$. Y is another continuous random variable with probability distribution function $F_Y(y) := P(Y \leq y)$. Variables X and Y are independent. Let $Z := \min(X, Y)$. What is the probability distribution function F_Z for Z ?

Solution

We have,

$$F_X(x) := P(X \leq x) \text{ and } F_Y(y) := P(Y \leq y)$$

Now,

$$F_Z(z) := P(Z \leq z) \dots (1)$$

So,

$$\begin{aligned} P(Z \leq z) &= 1 - P(Z > z) \\ &= 1 - P(\min(X, Y) > z) \\ &= 1 - P(X > z \text{ and } Y > z) \\ &= 1 - P(X > z)P(Y > z) \quad \because X \text{ and } Y \text{ are independent, } P(X \text{ and } Y) = P(X)P(Y) \\ &= 1 - (1 - P(X \leq z))(1 - P(Y \leq z)) \\ &= 1 - (1 - F_X(z))(1 - F_Y(z)) \\ &= 1 - (1 - F_X(z) - F_Y(z) + F_X(z)F_Y(z)) \\ &= F_X(z) + F_Y(z) - F_X(z)F_Y(z) \end{aligned}$$

Hence,

$$F_Z = F_X(z) + F_Y(z) - F_X(z)F_Y(z).$$

Question 5

[10 points] Let $Q : [n] \rightarrow [n]$ be a bijection. There is a standard technique to decompose Q as a product of disjoint cycles. For example, $C_1 := (i_1 := 1; Q(i_1); Q^2(i_1); \dots)$ gives the first cycle. Next, consider an element $i_2 \notin C_1$ to define $C_2 := (i_2; Q(i_2); Q^2(i_2); \dots)$ as the second cycle; and so on till all the elements $[n]$ are covered. Then, $(C_1; C_2; \dots; C_r)$ is essentially a unique cycle decomposition for Q .

Pick a *random* bijection Q on $[n]$. What is the expected number of cycles r for Q ?

Solution

We define a random variable A such that,

A = total number of cycles in Q

Now let L_i = length of cycle containing i

Let $A_i = L_i^{-1}$

Suppose first cycle has k numbers (i.e. $L_i = k \forall i \in \text{first cycle}$). So when we sum k^{-1} for every i (i.e. k^{-1} is added k times, it gives 1, i.e. first cycle). Adding in the similar way for all numbers, we get the total number of cycles. Hence, we get

$$A = \sum_{j=1}^n A_j$$

$$E[A] = E\left[\sum_{j=1}^n A_j\right]$$

Using linearity of expectation,

$$E[A] = \sum_{j=1}^n E[A_j] \quad \dots(1)$$

Using definition of expectation,

$$E[A_j] = \sum_{i=1}^n \frac{1}{i} P(A_j = \frac{1}{i}) \quad \dots(2)$$

Now we need to calculate $P(A_j = \frac{1}{i})$, i.e. probability that length of cycle containing j is i .

Counting favourable bijections:

Choose $i - 1$ numbers from the $n - 1$ numbers in Range (since j cannot be mapped to itself unless $i = 1$). This forms a cycle of i numbers. Now, number of different possible mappings among these i numbers are $(i - 1)!$. This is because j can be mapped to $i - 1$ values among these. Then, $Q(j)$ can be mapped to left $(i - 2)$ values and so on, giving value as $(i - 1)!$.

And then the rest of the numbers which are not present in this cycle (i.e. $n - i$) form bijection with the left $n - i$ numbers in Range using $(n - i)!$ ways. Hence,

Number of favourable bijections in which $(A_j = i^{-1}) = \binom{n-1}{i-1} (n - i)! (i - 1)!$

Total possible bijections = $n!$

So,

$$\begin{aligned} P(A_j = i^{-1}) &= \frac{\binom{n-1}{i-1} (n - i)! (i - 1)!}{n!} \\ P(A_j = i^{-1}) &= \frac{(n - i)!}{n!} \\ P(A_j = i^{-1}) &= \frac{1}{n} \quad \dots (3) \end{aligned}$$

Put (3) in (2),

$$\begin{aligned} E[A_j] &= \sum_{i=1}^n \frac{1}{i} \frac{1}{n} \\ E[A_j] &= \frac{1}{n} \sum_{i=1}^n \frac{1}{i} \\ E[A] &= \sum_{j=1}^n \frac{1}{n} \sum_{i=1}^n \frac{1}{i} \quad \text{From (1)} \\ E[A] &= \sum_{i=1}^n \frac{1}{i} \\ E[A] &\approx \ln(n) + \gamma \end{aligned}$$

Hence,

$$E[A] \approx \ln(n) + \gamma \quad (\text{where } \gamma \text{ is euler's constant})$$