CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair — Discussion 14

1 Short Tree Proofs

Note 5

Let $G = (V, E)$ be an undirected graph with $|V| \geq 1$.

(a) Prove that every connected component in an acyclic graph is a tree.

(b) Suppose G has k connected components. Prove that if G is acyclic, then $|E| = |V| - k$.

(c) Prove that a graph with $|V|$ edges contains a cycle.

Solution:

(a) Every connected component is connected, and acyclic because the graph is acyclic; by definition, this is a tree.

(b) Because each connected component is a tree, each connected component has $|V_i| - 1$ edges. The total number of edges is thus $\sum_i (|V_i| - 1) = |V| - k$.

(c) An acyclic graph has $|V| - k$ edges which cannot equal $|V|$, thus if a graph has $|V|$ edges it has a cycle.

2 Secret Sharing Practice

Consider the following secret sharing schemes and solve for asked variables.

(a) Suppose there is a bag of candy locked with a passcode between 0 and an integer n. Create a scheme for 5 trick-or-treaters such that they can only open the bag of candy if 3 of them agree to open it.

(b) Create a scheme for the following situation: There are 4 cats and 3 dogs in the neighborhood, and you want them to only be able to get the treats if the majority of the animals of each type are hungry. The treats are locked by a passcode between 0 and an integer n.

Solution:

(a) Solutions vary. The polynomial should be degree 2 and each trick-or-treater should be given the polynomial evaluated at one point.

(b) The guiding principle in this solution is that a polynomial of degree d, is uniquely determined by $d+1$ points. Let there be three polynomials, one for cats c, one for dogs d, and one joint one j that has the secret that actually unlocks the treats. c will be degree 2 since you need 3 cats to agree to get the 3 points to uniquely determine it. and d with be degree 1 since you need 2 dogs to agree to get the 2 points to uniquely determine it. The j will be degree 1 and $c(0)$ will be $j(1)$, and the $d(0)$ will be $j(2)$. This way you need both the point from the dogs and the point from the cats to uniquely determine j and otherwise you will be unable to determine the $j(0)$. This is also why we make $j(0)$ our secret.

3 Counting Subsets

Note 11 Consider the set S of all (possibly infinite) subsets of $\mathbb{N}$.

(a) Show that there is a bijection between S and $T = \{f : \mathbb{N} \to \{0,1\}\}$ (the set of all functions that map each natural number to 0 or 1).

(b) Prove or disprove: S is countable.

(c) Say that a function $f : \mathbb{N} \to \{0,1\}$ has *finite support* if it is non-zero on only a finite set of inputs. Let F denote the set of functions $f : \mathbb{N} \to \{0,1\}$ with finite support.

Prove that F is countably infinite.

Solution:

(a) Let $X \subseteq \mathbb{N}$. Define $f(x) = 1$ if $x \in X$, and 0 otherwise.

This is onto, since for a function f, the set that maps to it is $\{x \in \mathbb{N} \mid f(x) = 1\}$.

This is one-to-one. By showing the contrapositive, if two different set X and X' differ on some value x. Let $x \in X$ and $x \notin X'$. The function f and f' which they map to respectively will differ at $f(x) = 1$ and $f'(x) = 0$. Thus $f \neq f'$.

(b) Uncountable. Note that such f can be viewed as a binary encoding of a real number between 0 and 1, which exhibits a surjection from V to $[0,1]$.

(c) We give a bijection between F and $\mathbb{N}$. We encode an $f \in S$ as a binary number y_f, with the ith position (with $i = 0$ being the least significant digit) set to 1 if $f(i) = 1$. Note that this encoding always has finite length, excluding leading zeros, since the maximum i for which $f(i) = 1$ is finite. Thus the encoding always results in a natural number encoded in binary. This conversion is one-to-one, since each f and f' in F differ on at least one input and therefore y_f and $y_{f'}$ differ in at least one position. The conversion is onto, since every binary number represents a function with finite support. Since the natural numbers are countably infinite, and we have a bijection between F to $\mathbb{N}$, F is countably infinite.

An alternative bijection is between F and the subset of $\mathbb{N}$ that contains only numbers that are the product of distinct primes. Let $\{p_0 = 2, p_1 = 3, p_2 = 5, \dots\}$ be the set of all primes where

p_i is the $(i+1)$th prime. As shown in a previous homework, this set is infinite. Now consider a function $f(x)$ that is 1 exactly on inputs $x_1, x_2, \ldots, x_k$. Encode $f(x)$ as the natural number $p_{x_1} \times p_{x_2} \times \cdots \times p_{x_k}$. In other words, the function f is encoded as the natural number $2^{f(0)} \times 3^{f(1)} \times 5^{f(2)} \times 7^{f(3)} \times 11^{f(4)} \times \cdots$. This encoding is one-to-one, since the prime factorization of a number is unique. The encoding is onto, since every natural number that is composed of distinct primes corresponds to a function in F. Thus this is a bijection, and F is countable.

4 Strings

Note 10 How many different strings of length 5 only contain A, B, C? And how many such strings contain at least one of each characters?

Solution: The number of different strings of length 5 is 3^5 since each position have 3 different choices.

Let E_A be the set of strings that the character A is not used in the string. We define E_B, E_C similarly. Then the total number of "bad" strings is $|E_A \cup E_B \cup E_C|$.

By the Principle of Inclusion and Exclusion,

$$|E_A \cup E_B \cup E_C| = |E_A| + |E_B| + |E_C| - |E_A \cap E_B| - |E_A \cap E_C| - |E_B \cap E_C| + |E_A \cap E_B \cap E_C| = 3 \cdot 2^5 - 3 \cdot 1 = 93$$

where $|E_A \cap E_B| = |E_B \cap E_C| = |E_C \cap E_A| = 1$, and $|E_A \cap E_B \cap E_C| = 0$. Thus, the total number of valid string is $3^5 - 93 = 150$

5 Mario's Coins

Note 14 Mario owns three identical-looking coins. One coin shows heads with probability $1/4$, another shows heads with probability $1/2$, and the last shows heads with probability $3/4$.

(a) Mario randomly picks a coin and flips it. He then picks one of the other two coins and flips it. Let X_1 and X_2 be the events of the 1st and 2nd flips showing heads, respectively. Are X_1 and X_2 independent? Please prove your answer.

(b) Mario randomly picks a single coin and flips it twice. Let Y_1 and Y_2 be the events of the 1st and 2nd flips showing heads, respectively. Are Y_1 and Y_2 independent? Please prove your answer.

(c) Mario arranges his three coins in a row. He flips the coin on the left, which shows heads. He then flips the coin in the middle, which shows heads. Finally, he flips the coin on the right. What is the probability that it also shows heads?

Solution:

(a) X_1 and X_2 are not independent. Intuitively, the fact that X_1 happened gives some information about the first coin that was chosen; this provides some information about the second coin that was chosen (since the first and second coins can't be the same coin), which directly affects whether X_2 happens or not.

To make this formal, we compute

$$\mathbb{P}[X_1] = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{3}{4}\right) = \frac{1}{2}$$

By symmetry, $\mathbb{P}[X_2] = \mathbb{P}[X_1]$, so

$$\mathbb{P}[X_1]\,\mathbb{P}[X_2] = \frac{1}{4}.$$

But if we consider the probability that both X_1 and X_2 happen, we have

$$\begin{aligned}\mathbb{P}[X_1 \cap X_2] &= \frac{1}{6}\Big[\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) + \\ &\qquad \left(\frac{1}{2}\right)\left(\frac{3}{4}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{2}\right)\Big] \\ &= \frac{22}{96} = \frac{11}{48}\end{aligned}$$

which is not equal to $1/4$, violating the definition of independence.

(b) Y_1 and Y_2 are not independent. Intuitively, the fact that Y_1 happens gives some information about the coin that was picked, which directly influences whether Y_2 happens or not.

To make this formal, we compute

$$\mathbb{P}[Y_1] = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{3}{4}\right) = \frac{1}{2}$$

By symmetry, $\mathbb{P}[Y_2] = \mathbb{P}[Y_1]$, so

$$\mathbb{P}[Y_1]\,\mathbb{P}[Y_2] = \frac{1}{4}$$

But if we consider the probability that both Y_1 and Y_2 happen, we have

$$\mathbb{P}[Y_1 \cap Y_2] = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right)^2 + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right)^2 + \left(\frac{1}{3}\right)\left(\frac{3}{4}\right)^2 = \frac{14}{48} = \frac{7}{24}$$

which is not equal to $1/4$, violating the definition of independence.

(c) Let A be the coin with bias $1/4$, B be the fair coin, and C be the coin with bias $3/4$. There are six orderings, each with probability $1/6$: ABC, ACB, BAC, BCA, CAB, and CBA. Thus

$$\begin{aligned}&\mathbb{P}[\text{Third coin shows heads} \mid \text{First two coins show heads}] \\ &= \frac{\mathbb{P}[\text{All three coins show heads}]}{\mathbb{P}[\text{First two coins show heads}]} \\ &= \frac{(\frac{1}{4})(\frac{1}{2})(\frac{3}{4})}{11/48} \\ &= \frac{3/32}{11/48} = \frac{9}{22}.\end{aligned}$$

Note that the denominator was the probability calculated in part a, so we just plugged it in as $\frac{11}{48}$.

6 Sum of Poisson Variables

Note 19

Assume that you were given two independent Poisson random variables X_1, X_2. Assume that the first has mean λ_1 and the second has mean λ_2. Prove that $X_1 + X_2$ is a Poisson random variable with mean $\lambda_1 + \lambda_2$.

Hint: Recall the binomial theorem.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Solution:

To show that $X_1 + X_2$ is a Poisson random variable with mean $\lambda_1 + \lambda_2$, we have show that

$$\mathbb{P}[(X_1 + X_2) = i] = \frac{(\lambda_1 + \lambda_2)^i}{i!} \mathrm{e}^{-(\lambda_1+\lambda_2)}.$$

We proceed as follows:

$$\begin{aligned}
\mathbb{P}[(X_1 + X_2) = i] &= \sum_{k=0}^{i} \mathbb{P}[X_1 = k, X_2 = (i-k)] = \sum_{k=0}^{i} \frac{\lambda_1^k}{k!} \mathrm{e}^{-\lambda_1} \cdot \frac{\lambda_2^{i-k}}{(i-k)!} \mathrm{e}^{-\lambda_2} \\
&= \mathrm{e}^{-\lambda_1} \mathrm{e}^{-\lambda_2} \sum_{k=0}^{i} \frac{1}{k!(i-k)!} \lambda_1^k \lambda_2^{i-k} = \frac{\mathrm{e}^{-\lambda_1} \mathrm{e}^{-\lambda_2}}{i!} \sum_{k=0}^{i} \frac{i!}{k!(i-k)!} \lambda_1^k \lambda_2^{i-k} \\
&= \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{i!} \sum_{k=0}^{i} \binom{i}{k} \lambda_1^k \lambda_2^{i-k} = \frac{\mathrm{e}^{-(\lambda_1+\lambda_2)}}{i!} (\lambda_1 + \lambda_2)^i
\end{aligned}$$

In the last line, we use the binomial expansion.

7 Balls in Bins

Note 15

You are throwing k balls into n bins. Let X_i be the number of balls thrown into bin i.

(a) What is $\mathbb{E}[X_i]$?

(b) What is the expected number of empty bins?

(c) Define a collision to occur when a ball lands in a nonempty bin (if there are n balls in a bin, count that as $n-1$ collisions). What is the expected number of collisions?

Solution:

(a) We will use linearity of expectation. Note that the expectation of an indicator variable is just the probability the indicator variable = 1. (Verify for yourself that is true).

$$\begin{aligned}\mathbb{E}[X_i] &= \mathbb{P}[\text{ball 1 falls into bin } i] + \mathbb{P}[\text{ball 2 falls into bin } i] + \cdots + \mathbb{P}[\text{ball } k \text{ falls into bin } i] \\ &= \frac{1}{n} + \cdots + \frac{1}{n} = \frac{k}{n}.\end{aligned}$$

(b) Let I_i be the indicator variable denoting whether bin i ends up empty. This can happen if and only if all the balls end in the remaining $n-1$ bins, and this happens with a probability of $\left(\frac{n-1}{n}\right)^k$. Hence the expected number of empty bins is

$$\mathbb{E}[I_1 + \ldots + I_n] = \mathbb{E}[I_1] + \ldots + \mathbb{E}[I_n] = n\left(\frac{n-1}{n}\right)^k$$

(c) The number of collisions is the number of balls minus the number of occupied bins, since the first ball of every occupied bin is not a collision.

$$\begin{aligned}\mathbb{E}[\text{collisions}] &= k - \mathbb{E}[\text{occupied bins}] = k - n + \mathbb{E}[\text{empty locations}] \\ &= k - n + n\left(1 - \frac{1}{n}\right)^k\end{aligned}$$

8 Inequality Practice

Note 17

(a) X is a random variable such that $X \geq -5$ and $\mathbb{E}[X] = -3$. Find an upper bound for the probability of X being greater than or equal to -1.

(b) Y is a random variable such that $Y \leq 10$ and $\mathbb{E}[Y] = 1$. Find an upper bound for the probability of Y being less than or equal to -1.

(c) You roll a die 100 times. Let Z be the sum of the numbers that appear on the die throughout the 100 rolls. Compute $\text{Var}(Z)$. Then use Chebyshev's inequality to bound the probability of the sum Z being greater than 400 or less than 300.

Solution:

(a) We want to use Markov's Inequality, but recall that Markov's Inequality only works with non-negative random variables. So, we define a new random variable $\tilde{X} = X + 5$, where $\tilde{X}$ is always non-negative, so we can use Markov's on $\tilde{X}$. By linearity of expectation, $\mathbb{E}[\tilde{X}] = -3 + 5 = 2$. So, $\mathbb{P}[\tilde{X} \geq 4] \leq 2/4 = 1/2$.

(b) We again use Markov's Inequality. Similarly, define $\tilde{Y} = -Y + 10$, and $\mathbb{E}[\tilde{Y}] = -1 + 10 = 9$. $P[Y \leq -1] = P[-Y \geq 1] = P[-Y + 10 \geq 11] \leq 9/11$.

(c) Let Z_i be the number on the die for the ith roll, for $i = 1, \ldots, 100$. Then, $Z = \sum_{i=1}^{100} Z_i$. By linearity of expectation, $\mathbb{E}[Z] = \sum_{i=1}^{100} \mathbb{E}[Z_i]$.

$$\mathbb{E}[Z_i] = \sum_{j=1}^{6} j \cdot \mathbb{P}[Z_i = j] = \sum_{j=1}^{6} j \cdot \frac{1}{6} = \frac{1}{6} \cdot \sum_{j=1}^{6} j = \frac{1}{6} \cdot 21 = \frac{7}{2}$$

Then, we have $\mathbb{E}[Z] = 100 \cdot (7/2) = 350$.

$$\mathbb{E}[Z_i^2] = \sum_{j=1}^{6} j^2 \cdot \mathbb{P}[Z_i = j] = \sum_{j=1}^{6} j^2 \cdot \frac{1}{6} = \frac{1}{6} \cdot \sum_{j=1}^{6} j^2 = \frac{1}{6} \cdot 91 = \frac{91}{6}$$

Then, we have

$$\mathrm{Var}(Z_i) = \mathbb{E}[Z_i^2] - \mathbb{E}[Z_i]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12},$$

Since the Z_is are independent, and therefore uncorrelated, we can add the $\mathrm{Var}(Z_i)$s to get $\mathrm{Var}(Z) = 100(35/12)$.

Finally, we note that we can upper bound $\mathbb{P}[|Z - 350| > 50]$ with $\mathbb{P}[|Z - 350| \geq 50]$.

Putting it all together, we use Chebyshev's to get

$$\mathbb{P}[|Z - 350| > 50] < \mathbb{P}[|Z - 350| \geq 50] \leq \frac{100(35/12)}{50^2} = \frac{7}{60}.$$

9 Exponential Distributions: Lightbulbs

Note 21

A brand new lightbulb has just been installed in our classroom, and you know the life span of a lightbulb is exponentially distributed with a mean of 50 days.

(a) Suppose an electrician is scheduled to check on the lightbulb in 30 days and replace it if it is broken. What is the probability that the electrician will find the bulb broken?

(b) Suppose the electrician finds the bulb broken and replaces it with a new one. What is the probability that the new bulb will last at least 30 days?

(c) Suppose the electrician finds the bulb in working condition and leaves. What is the probability that the bulb will last at least another 30 days?

Solution:

(a) Let $X \sim \text{Exponential}(1/50)$ be the time until the bulb is broken. For an exponential random variable with parameter λ, the density function is $f_X(x) = \lambda \, \mathrm{e}^{-\lambda x}$ for $x > 0$. So in this case

$\lambda = 1/50$. Thus we can integrate the density to find the probability that the lightbulb broke in the first 30 days:

$$\mathbb{P}[X < 30] = \int_0^{30} \left(\frac{1}{50} \cdot \mathrm{e}^{-x/50}\right) \mathrm{d}x = 1 - \mathrm{e}^{-30/50} = 1 - \mathrm{e}^{-3/5} \approx 0.451.$$

(b) The new bulb's waiting time Y is i.i.d. with the old bulb's. So the answer is

$$\mathbb{P}[Y > 30] = 1 - \mathbb{P}[Y < 30] = 1 - (1 - \mathrm{e}^{-3/5}) = \mathrm{e}^{-3/5} \approx 0.549.$$

(c) The bulb is memoryless, so the probability it will last 60 days given that it has lasted 30 days, is just the probability it will last 30 days:

$$\mathbb{P}[X > 60 \mid X > 30] = \mathbb{P}[X - 30 > 30 \mid X > 30] = \mathbb{P}[X > 30] = \mathrm{e}^{-3/5} \approx 0.549.$$

10 Continuous Probability Continued

Note 17

For the following questions, please briefly justify your answers or show your work.

(a) Assume $\mathrm{Bob}_1, \mathrm{Bob}_2, \ldots, \mathrm{Bob}_k$ each hold a fair coin whose two sides show numbers instead of heads and tails, with the numbers on Bob_i's coin being i and $-i$. Each Bob tosses their coin n times and sums up the numbers he sees; let's call this number X_i. For large n, what is the distribution of $(X_1 + \cdots + X_k)/\sqrt{n}$ approximately equal to?

(b) If $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables of mean μ and variance σ^2, what is $\lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k - \mu}{\sigma n^{\alpha}} \in [-1, 1]\right]$ for $\alpha \in [0, 1]$ (your answer may depend on α and Φ, the CDF of a $N(0,1)$ variable)?

Solution:

(a) $\boxed{N\left(0, \sum_{i=1}^{k} i^2\right)}$.

$(X_1 + \cdots + X_k)/\sqrt{n} = \frac{X_1}{\sqrt{n}} + \cdots + \frac{X_k}{\sqrt{n}}$, and since each $\frac{X_i}{\sqrt{n}}$ converges to $N\left(0, i^2\right)$ by the central limit theorem, their sum must converge to $N\left(0, \sum_{i=1}^{k} i^2\right)$. Alternatively, if we let X_j^i be the j^{th} coin toss of Bob_i, then $(X_1 + \cdots + X_k)/\sqrt{n} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (X_j^1 + \cdots + X_j^k)$. But the $Y_j = X_j^1 + \ldots X_j^k$ themselves are i.i.d. variables of mean 0 and variance $\sum_{i=1}^{k} i^2$, and so the central limit theorem again implies a limiting distribution of $N\left(0, \sum_{i=1}^{k} i^2\right)$ (this constitutes an alternative proof of the fact that the sum of Gaussians is also a Gaussian, which we showed in class).

(b) $\lim_{n\to\infty} \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k - \mu}{\sigma n^\alpha} \in [-1,1]\right] = \begin{cases} 1, & \text{if } \alpha > \frac{1}{2}, \\ \Phi(1) - \Phi(-1), & \text{if } \alpha = \frac{1}{2}, \\ 0, & \text{if } \alpha < \frac{1}{2} \end{cases}$.

For $\alpha > \frac{1}{2}$, the reasoning is exactly as in the law of large numbers: By Chebyshev's inequality, we have $1 - \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma n^\alpha} \in [-1,1]\right] = \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma n^\alpha} \notin [-1,1]\right] \le \frac{1}{n^{2\alpha-1}} \xrightarrow{n\to\infty} 0$. The $\alpha = \frac{1}{2}$ case is a direct consequence of the central limit theorem, while the $\alpha < \frac{1}{2}$ case follows indirectly from it: $\mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma n^\alpha} \in [-1,1]\right] = \mathbb{P}\left[\sum_{k=1}^{n} \frac{X_k-\mu}{\sigma\sqrt{n}} \in \left[-\frac{1}{n^{\frac{1}{2}-\alpha}}, \frac{1}{n^{\frac{1}{2}-\alpha}}\right]\right]$
$\approx \mathbb{P}\left[N(0,1) \in \left[-\frac{1}{n^{\frac{1}{2}-\alpha}}, \frac{1}{n^{\frac{1}{2}-\alpha}}\right]\right] \xrightarrow{n\to\infty} 0$.

11 Three Tails

Note 22

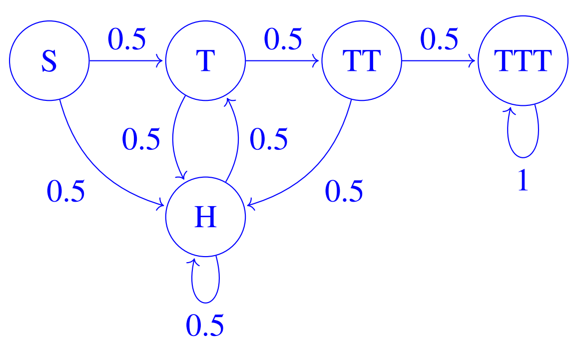
You flip a fair coin until you see three tails in a row. What is the average number of heads that you'll see until getting TTT?

Hint: How is this different than the number of *coins* flipped until getting TTT?

Solution:

We can model this problem as a Markov chain with the following states:

- S: Start state, which we are only in before flipping any coins.
- H: We see a head, which means no streak of tails currently exists.
- T: We've seen exactly one tail in a row so far.
- TT: We've seen exactly two tails in a row so far.
- TTT: We've accomplished our goal of seeing three tails in a row and stop flipping.



We can write the first step equations and solve for $\beta(S)$, only counting heads that we see since we are not looking for the total number of flips. The equations are as follows:

$$\beta(S) = 0.5\beta(T) + 0.5\beta(H) \tag{1}$$
$$\beta(H) = 1 + 0.5\beta(H) + 0.5\beta(T) \tag{2}$$
$$\beta(T) = 0.5\beta(TT) + 0.5\beta(H) \tag{3}$$
$$\beta(TT) = 0.5\beta(H) + 0.5\beta(TTT) \tag{4}$$
$$\beta(TTT) = 0 \tag{5}$$

From equation (2), we see that
$$0.5\beta(H) = 1 + 0.5\beta(T)$$
and can substitute that into equation (3) to get
$$0.5\beta(T) = 0.5\beta(TT) + 1.$$

Substituting this into equation (4), we can deduce that $\beta(TT) = 4$. This allows us to conclude that $\beta(T) = 6$, $\beta(H) = 8$, and $\beta(S) = 7$. On average, we expect to see 7 heads before flipping three tails in a row.