

1 Cliques in Random Graphs

Consider the graph $G = (V, E)$ on n vertices which is generated by the following random process: for each pair of vertices u and v , we flip a fair coin and place an (undirected) edge between u and v if and only if the coin comes up heads. Now consider:

- What is the size of the sample space?
- A k -clique in graph is a set S of k vertices which are pairwise adjacent (every pair of vertices is connected by an edge). For example a 3-clique is a triangle. Let's call the event that S forms a clique E_S . What is the probability of E_S for a particular set S of k vertices?
- For two sets of vertices $V_1 = \{v_1, \dots, v_\ell\}$ and $V_2 = \{w_1, \dots, w_k\}$, are E_{V_1} and E_{V_2} independent?
- Prove that $\binom{n}{k} \leq n^k$.
- Prove that the probability that the graph contains a k -clique, for $k \geq 4 \log n + 1$, is at most $1/n$. (The log is taken base 2). *Hint:* Apply the union bound and part (d).

Solution:

- Between every pair of vertices, there is either an edge or not. Since there are two choices for each of the $\binom{n}{2}$ pairs of vertices, the size of the sample space is $2^{\binom{n}{2}}$.
- For a fixed set of k vertices to be a k -clique, all of the $\binom{k}{2}$ pairs of those vertices have to be connected by an edge. The probability of this event is $1/2^{\binom{k}{2}}$.
- E_{V_1} and E_{V_2} are independent if and only if V_1 and V_2 share at most one vertex: If V_1 and V_2 share at most one vertex, then since edges are added independently of each other, we have

$$\mathbb{P}(E_{V_1} \cap E_{V_2}) = \mathbb{P}(\text{all edges in } V_1 \text{ and all edges in } V_2 \text{ are present}) = \left(\frac{1}{2}\right)^{\binom{|V_1|}{2}} \cdot \left(\frac{1}{2}\right)^{\binom{|V_2|}{2}} = \mathbb{P}(E_{V_1}) \cdot \mathbb{P}(E_{V_2}).$$

Conversely, if V_1 and V_2 share at least two vertices, then their intersection $V_3 = V_1 \cap V_2$ has at least 2 elements, and whence

$$\mathbb{P}(E_{V_1} \cap E_{V_2}) = \left(\frac{1}{2}\right)^{\binom{|V_3|}{2}} \cdot \left(\frac{1}{2}\right)^{\binom{|V_1|}{2} - \binom{|V_3|}{2}} \cdot \left(\frac{1}{2}\right)^{\binom{|V_2|}{2} - \binom{|V_3|}{2}} = \left(\frac{1}{2}\right)^{\binom{|V_1|}{2} + \binom{|V_2|}{2} - \binom{|V_3|}{2}} \neq \mathbb{P}(E_{V_1}) \cdot \mathbb{P}(E_{V_2}).$$

(d) The algebraic solution is an application of the definition of $\binom{n}{k}$:

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \quad (1)$$

$$\leq n \cdot (n-1) \cdots (n-k+1) \quad (2)$$

$$\leq n^k \quad (3)$$

(e) Let A_S denote the event that S is a k -clique, where $S \subseteq V$ is of size k . Then, the event that the graph contains a k -clique can be described as the union of A_S 's over all $S \subseteq V$ of size k . Using the union bound,

$$\mathbb{P} \left[\bigcup_{S \subseteq V, |S|=k} A_S \right] \leq \sum_{S \subseteq V, |S|=k} \mathbb{P}[A_S] = \sum_{S \subseteq V, |S|=k} \frac{1}{2^{\binom{k}{2}}}.$$

Now, since there are $\binom{n}{k}$ ways of choosing a subset $S \subseteq V$ of size k , the right-hand side of the above equality is

$$\frac{\binom{n}{k}}{2^{\binom{k}{2}}} = \frac{\binom{n}{k}}{2^{k(k-1)/2}} \leq \frac{n^k}{(2^{(k-1)/2})^k} \leq \frac{n^k}{(2^{(4 \log n + 1 - 1)/2})^k} = \frac{n^k}{(2^{2 \log n})^k} = \frac{n^k}{n^{2k}} = \frac{1}{n^k} \leq \frac{1}{n}.$$

2 Balls and Bins, All Day Every Day

Suppose n balls are thrown into n labeled bins one at a time, where n is a positive *even* integer.

- What is the probability that exactly k balls land in the first bin, where k is an integer $0 \leq k \leq n$?
- What is the probability p that at least half of the balls land in the first bin? (You may leave your answer as a summation.)
- Using the union bound, give a simple upper bound, in terms of p , on the probability that some bin contains at least half of the balls.
- What is the probability, in terms of p , that at least half of the balls land in the first bin, or at least half of the balls land in the second bin?
- After you throw the balls into the bins, you walk over to the bin which contains the first ball you threw, and you randomly pick a ball from this bin. What is the probability that you pick up the first ball you threw? (Again, leave your answer as a summation.)

Solution:

- The probability that a particular ball lands in the first bin is $1/n$. We need exactly k balls to land in the first bin, which occurs with probability $(1/n)^k$, and we need exactly $n-k$ balls to land in a different bin, which occurs with probability $(1-1/n)^{n-k}$, and there are $\binom{n}{k}$ ways to choose which of the k balls land in first bin. Thus, the probability is $\binom{n}{k} (1/n)^k (1-1/n)^{n-k}$.

- (b) This is the summation over $k = n/2, \dots, n$ of the probabilities computed in the first part, i.e., $\sum_{k=n/2}^n \binom{n}{k} (1/n)^k (1 - 1/n)^{n-k}$.
- (c) The event that some bin has at least half of the balls is the union of the events $A_k, k = 1, \dots, n$, where A_k is the event that bin k has at least half of the balls. By the union bound, $\mathbb{P}(\bigcup_{k=1}^n A_k) \leq \sum_{k=1}^n \mathbb{P}(A_k) = np$.
- (d) The probability that the first bin has at least half of the balls is p ; similarly, the probability that the second bin has at least half of the balls is also p . There is overlap between these two events, however: the first bin has half of the balls and the second bin has the second half of the balls. The probability of this event is $\binom{n}{n/2} n^{-n}$: there are n^n total possible configurations for the n balls to land in the bins, but if we require exactly $n/2$ of the balls to land in the first bin and the remaining balls to land in the second bin, there are $\binom{n}{n/2}$ ways to choose which balls land in the first bin. By the principle of inclusion-exclusion, our desired probability is $p + p - \binom{n}{n/2} n^{-n} = 2p - \binom{n}{n/2} n^{-n}$.
- (e) Condition on the number of balls in the bin. First we calculate the probability $\mathbb{P}(A_k)$, where A_k is the event that, in addition to the first ball you threw, an additional $k - 1$ of the other $n - 1$ balls landed in this bin, which by the reasoning in Part (a) has probability

$$\mathbb{P}(A_k) = \binom{n-1}{k-1} (1/n)^{k-1} (1 - 1/n)^{n-k} .$$

If we let B be the event that we pick up the first ball we threw, then

$$\mathbb{P}(B | A_k) = 1/k$$

since we are equally likely to pick any of the k balls in the bin. Thus the overall probability we are looking for is, by an application of the law of total probability,

$$\mathbb{P}(B) = \sum_{k=1}^n \mathbb{P}(A_k \cap B) = \sum_{k=1}^n \mathbb{P}(A_k) \mathbb{P}(B | A_k) = \sum_{k=1}^n \frac{1}{k} \binom{n-1}{k-1} \left(\frac{1}{n}\right)^{k-1} \left(1 - \frac{1}{n}\right)^{n-k} .$$

3 Lie Detector

A lie detector is known to be $4/5$ reliable when the person is guilty and $9/10$ reliable when the person is innocent. If a suspect is chosen from a group of suspects of which only $1/100$ have ever committed a crime, and the test indicates that the person is guilty, what is the probability that he is guilty?

Solution:

Let A denote the event that the test indicates that the person is guilty, and B the event that the person is actually guilty. Note that

$$\mathbb{P}[B] = \frac{1}{100}, \quad \mathbb{P}[\bar{B}] = \frac{99}{100}, \quad \mathbb{P}[A | B] = \frac{4}{5}, \quad \mathbb{P}[A | \bar{B}] = \frac{1}{10}.$$

By Bayes' Rule and the Total Probability Rule the desired probability is

$$\mathbb{P}[B | A] = \frac{\mathbb{P}[B]\mathbb{P}[A | B]}{\mathbb{P}[A]} = \frac{\mathbb{P}[B]\mathbb{P}[A | B]}{\mathbb{P}[B]\mathbb{P}[A | B] + \mathbb{P}[\bar{B}]\mathbb{P}[A | \bar{B}]} = \frac{(1/100)(4/5)}{(1/100)(4/5) + (99/100)(1/10)} = \frac{8}{107}$$

4 Flipping Coins

Consider the following scenarios, where we apply probability to a game of flipping coins. In the game, we flip one coin each round. The game will not stop until two consecutive heads appear.

- What is the probability that the game ends by flipping exactly five coins?
- Given that the game ends after flipping the fifth coin, what is the probability that three heads appear in the sequence?
- If we change the rule that the game will not stop until three consecutive tails or three consecutive heads appear, what is the probability that the game stops by flipping at most six coins?

Solution:

- If the game ends by flipping exactly five coins, we know the flipping results of last three coins must be $\{T, H, H\}$. For the first two coins, the results can be $\{H, T\}$, $\{T, H\}$ or $\{T, T\}$. So the probability equals to $\frac{3}{4} \times \frac{1}{8} = \frac{3}{32}$.
- Given the condition that the game ends after flipping the fifth coin, the only possible sequences containing three heads are $\{H, T, T, H, H\}$ and $\{T, H, T, H, H\}$. Let A denote the event that the game ends after flipping the fifth coin, B denote the event that three heads appear when the game ends. Then $\mathbb{P}(A \cap B) = \frac{2}{32}$ and $\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{2}{3}$.
- We only consider the case that the game ends only because of three consecutive heads appearing. If the game ends in three rounds, the result must be $\{H, H, H\}$ and the probability is $\frac{1}{8}$. If the game ends in four rounds, the results must be $\{T, H, H, H\}$ and the probability is $\frac{1}{16}$. If the game ends in five rounds, the results must be $\{H, T, H, H, H\}$ or $\{T, T, H, H, H\}$ and the probability is $\frac{1}{16}$. If the game ends in six rounds, the last four coins must be $\{T, H, H, H\}$ and the first two coins cannot be $\{T, T\}$, so the probability is $\frac{3}{4} \times \frac{1}{16} = \frac{3}{64}$. The total probability is

$$\frac{1}{8} + 2 \times \frac{1}{16} + \frac{3}{64} = \frac{19}{64}$$

. By symmetry, the probability that the game ends in the sixth round because of three consecutive tails is also $\frac{19}{64}$. So the final answer is

$$2 \times \frac{19}{64} = \frac{19}{32}$$

5 To Be Fair

Suppose you have a biased coin with $\mathbb{P}(\text{heads}) \neq 0.5$. How could you use this coin to simulate a fair coin? (*Hint*: Think about pairs of tosses.)

Solution:

Let's think about the experiment of throwing the biased coin twice as hinted towards in the hint: Its sample space is $\Omega = \{\text{HH}, \text{TT}, \text{HT}, \text{TH}\}$ with corresponding probability function $\mathbb{P}(\text{HH}) = p^2, \mathbb{P}(\text{TT}) = (1-p)^2$ and $\mathbb{P}(\text{HT}) = \mathbb{P}(\text{TH}) = p(1-p)$. Neither of these probabilities is $1/2$ which is what would be required for simulating a fair coin. However, we can generate new probabilities by looking at conditional probabilities! In particular, knowing that $\mathbb{P}(\text{HT}) = \mathbb{P}(\text{TH})$ informs us that $\mathbb{P}(\text{HT} \mid \{\text{HT}, \text{TH}\}) = \frac{\mathbb{P}(\text{HT})}{\mathbb{P}(\{\text{HT}, \text{TH}\})} = \frac{\mathbb{P}(\text{TH})}{\mathbb{P}(\{\text{HT}, \text{TH}\})} = \mathbb{P}(\text{TH} \mid \{\text{HT}, \text{TH}\})$, and since $\mathbb{P}(\text{HT} \mid \{\text{HT}, \text{TH}\}) + \mathbb{P}(\text{TH} \mid \{\text{HT}, \text{TH}\}) = 1$, it must be true that

$$\mathbb{P}(\text{HT} \mid \{\text{HT}, \text{TH}\}) = \mathbb{P}(\text{TH} \mid \{\text{HT}, \text{TH}\}) = 1/2.$$

That is, to simulate a fair coin we can throw the biased coin twice until we observe either HT or TH, since the probability of either showing up first is exactly $1/2$.