

CS70 Homework 8

Michael Pham

Spring 2023

Contents

1	Probability Warm-Up	3
2	Five Up	4
3	Past Probabilified	6
4	Cliques in Random Graphs	8
5	PIE Extended	10
6	Independent Complements	12

1 Probability Warm-Up

Problem 1.1. Suppose that we have a bucket of 30 green balls and 70 orange balls. If we pick 15 balls uniformly out of the bucket, what is the probability of getting exactly k green balls (assuming $0 \leq k \leq 15$) if the sampling is done **with** replacement, i.e. after we take a ball out the bucket we return the ball back to the bucket for the next round?

Solution. The probability of getting k green balls out of 15 trials is:

$$P(k) = \binom{15}{k} \left(\frac{3}{10}\right)^k \left(\frac{7}{10}\right)^{15-k}$$

■

Problem 1.2. Same as the previous part, but the sampling is **without** replacement, i.e. after we take a ball out the bucket we **do not** return the ball back to the bucket.

Solution. If, instead, it was done without replacement, then we observe that the probability will be:

$$P(k) = \frac{\binom{30}{k} \binom{70}{15-k}}{\binom{100}{15}}$$

■

Problem 1.3. If we roll a regular, 6-sided die 5 times. What is the probability that at least one value is observed more than once?

Solution. In order to calculate the probability of this event, which we will denote by $P(A)$, we first calculate the probability of getting a roll where each number is distinct, and then find its complement.

To do this, observe that $|\Omega| = 6^5$. Next, we want to count the number of ways to get 5 distinct values from the 6 values on the dice, which is $5! \binom{6}{5}$. Thus, the probability of getting all distinct numbers from 5 rolls is:

$$\frac{5! \binom{6}{5}}{6^5} = \frac{5}{54}$$

And its complement, which is the probability we are interested in, is $1 - \frac{5}{54} = \frac{49}{54}$.

■

2 Five Up

Say you toss a coin five times, and record the outcomes. For the three questions below, you can assume that order matters in the outcome, and that the probability of heads is some p in $0 < p < 1$, but *not* that the coin is fair ($p = 0.5$).

Problem 2.1. What is the size of the sample space, $|\Omega|$?

Solution. The size of the sample space is $2^5 = 32$. ■

Problem 2.2. How many elements of Ω have exactly three heads?

Solution. There are $\binom{5}{3} = 10$ elements which have exactly three heads. ■

Problem 2.3. How many elements of Ω have three or more heads?

Solution. In order to find the number of elements which have three or more heads, we can instead find the number of elements that have no heads, one heads, and two heads first. From there, we can subtract it from $|\Omega|$:

1. For no heads, there's only 1 element.
2. For one head, there are only 5 elements.
3. For two heads, there are 10 elements.

Thus, the number of elements with less than three heads are 16. So, $P(H \geq 3) = 32 - 16 = 16$. ■

For the next three questions, you can assume that the coin is fair (i.e. heads comes up with $p = 0.5$, and tails otherwise).

Problem 2.4. What is the probability that you will observe the sequence HHHTT? What about HHHHT?

Solution. The probability of observing the sequence HHHTT is $\frac{1}{32}$. Similarly, the sequence HHHHT is also $\frac{1}{32}$. ■

Problem 2.5. What is the chance of observing at least one head?

Solution. The probability of observing at least one head is $1 - P(H = 0)$. In this case, we see that the probability of getting no heads is $\frac{1}{32}$, so $P(H \geq 1) = 1 - P(H = 0) = 1 - \frac{1}{32} = \frac{31}{32}$. ■

Problem 2.6. What about the chance of observing three or more heads?

Solution. The chances of observing three or more heads is simply $\frac{1}{2}$. ■

For the final three questions, you can instead assume the coin is biased so that it comes up heads with probability $p = \frac{2}{3}$.

Problem 2.7. What is the chance of observing the outcome HHHTT? What about HHHHT?

Solution. For HHHTT, the probability of observing such an outcome is $\frac{2(2)(2)(1)(1)}{3^5} = \frac{8}{243}$.
For a sequence of HHHHT, it is $\frac{16}{243}$. ■

Problem 2.8. What about the chance of at least one head?

Solution. The probability of getting no heads is $\left(\frac{1}{3}\right)^5 = \frac{1}{243}$. So, the probability of at least one head is $1 - \frac{1}{243} = \frac{242}{243}$. ■

Problem 2.9. What about the chance of ≥ 3 heads?

Solution. To calculate the probability of getting ≥ 3 heads, we can first find the following:

1. $P(H = 0) = \frac{1}{243}$
2. $P(H = 1) = \binom{5}{1} \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^4 = \frac{10}{243}$
3. $P(H = 2) = \binom{5}{2} \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^3 = \frac{40}{243}$

So, $P(H \geq 3) = 1 - \frac{1}{243} - \frac{10}{243} - \frac{40}{243} = \frac{192}{243}$. ■

3 Past Probabilified

In this question we review some of the past CS70 topics, and look at them probabilistically. For the following experiments,

- i. Define an appropriate sample space Ω .
- ii. Give the probability function $\Pr[\omega]$.
- iii. Compute $\Pr[E_1]$.
- iv. Compute $\Pr[E_2]$.

Problem 3.1. Fix a prime $q > 2$, and uniformly sample twice with replacement from $\{0, \dots, q-1\}$ (assume we have two $\{0, \dots, q-1\}$ -sided fair dice and we roll them). Then multiply these two numbers with each other modulo q .

Let E_1 = The resulting product is 0.

Let E_2 = The product is $(q-1)/2$.

Solution. First, we observe that the sample space $\Omega = \{(0, 0), (0, 1), \dots, (q-1, q-1)\}$.

Next, the probability function $\Pr[\omega] = \frac{|\omega|}{q^2}$.

Next, to find $\Pr[E_1]$, we observe that for the resulting product to be $0 \pmod{q}$, at least one of the numbers must be 0. There are $(q-1) + (q-1) + 1 = 2q-1$ pairs. Thus, we have that $\Pr[E_1] = \frac{2q-1}{q^2}$.

For $\Pr[E_2]$, we observe that for every number from $a \in \{1, \dots, q-1\}$, because q is prime then there exists a unique number equal to $a^{-1} \left(\frac{q-1}{2}\right) \pmod{q}$. Thus, we observe that there are $q-1$ of such pairs, yielding us $\Pr[E_2] = \frac{q-1}{q^2}$. ■

Problem 3.2. Make a graph on v vertices by sampling uniformly at random from all possible edges. Here, assume for each edge we flip a fair coin; if it comes up heads, we include the edge in the graph, and otherwise we exclude that edge from the graph.

Let E_1 = The graph is complete.

Let E_2 = vertex v_1 has degree d .

Solution. For the sample space, we observe that it's simply a set of all possible graphs with v vertices; in other words, $\Omega = \mathcal{P}(K_v)$, where K_v is the set of edges for a complete graph on v vertices.

Next, the probability function $\Pr[\omega] = \frac{|\omega|}{2^{\binom{v}{2}}}$.

To find $\Pr[E_1]$, we observe that there's only one graph such that this can occur; thus, we have $\Pr[E_1] = \frac{1}{2^{\binom{v}{2}}}$.

To find $\Pr[E_2]$, we observe that the number of graphs where a vertex v_1 has degree d is ■

Problem 3.3. Create a random stable matching instance by having each person's preference list be a uniformly random permutation of the opposite entity's list (make the preference list for each individual job and each individual candidate a random permutation of the opposite entity's list). Finally, create a uniformly random pairing by matching jobs and candidates up uniformly at random (note that in this pairing, (1) a candidate cannot be matched with two different jobs, and a job cannot be matched with two different candidates (2) the pairing does not have to be stable).

Let E_1 = All jobs have distinct favorite candidates.

Let E_2 = The resulting pairing is the candidate-optimal stable pairing.

Solution. To begin with, the sample space Ω is $\{\{\{c_1, \dots, c_n\}, \{j_1, \dots, j_n\}\}, \dots\}$, where the elements contain every permutation of the preference lists of each candidates/jobs.

Next, we see that $\Pr[\omega] = \frac{|\omega|}{(n!)^{2n} \frac{n!}{\frac{n}{2}! 2^{\frac{n}{2}}}}$

For $\Pr[E_1]$, it is $\frac{n!}{(n!)^{2n} \frac{n!}{\frac{n}{2}! 2^{\frac{n}{2}}}}$ ■

4 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.

Problem 4.1. What is the size of the sample space?

Solution. The size of the sample space, that is $|\Omega|$, is $2^{\binom{n}{2}}$. ■

Problem 4.2. A k -clique in a 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 E_S be the event that a set S forms a clique. What is the probability of E_S for a particular set S of k vertices?

Solution. We observe that the probability is that a set S of k vertices forms a clique is:

$$\Pr[E_S] = \frac{1}{2^{\binom{k}{2}}}$$

Problem 4.3. Suppose that $V_1 = \{v_1, \dots, v_\ell\}$ and $V_2 = \{w_1, \dots, w_k\}$ are two arbitrary sets of vertices. What conditions must V_1 and V_2 satisfy in order for E_{V_1} and E_{V_2} to be independent? Prove your answer.

Solution. For E_{V_1} and E_{V_2} to be independent, then $\Pr[E_{V_1} \cap E_{V_2}] = \Pr[E_{V_1}] \Pr[E_{V_2}]$. For this to be the case, we observe then that there can't be 2 or more vertices shared between the two probabilities.

We observe that if two vertices are shared between the two, and if one of the events happens to be true, the other event's probability will get doubled.

Thus, for it to be independent, they can't share vertices. ■

Problem 4.4. Prove that $\binom{n}{k} \leq n^k$.

Solution. We want to show that for all $k \in \mathbb{Z}$, we have that $\binom{n}{k} \leq n^k$.

First, we shall prove equality.

For $k = 0$, we observe that $\binom{n}{0} = 1 = n^0$.

Similarly, for $k = 1$, we see that $\binom{n}{1} = n = n^1$.

Now, for $k > 1$, we observe the following:

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} \\ &< \frac{n!}{(n-k)!} \\ &= (n)(n-1) \cdots (n-k+1) \end{aligned}$$

From here, we observe that there will be k factors for $n(n-1) \cdots (n-k+1)$, with at least one factor less than n . Meanwhile, for $k > 1$, $n^k = n(n) \cdots (n)$ has k factors where all of them are n . Thus, we can conclude the following for $k > 1$:

$$n(n-1) \cdots (n-k+1) < n^k$$

Therefore, we see that $\binom{n}{k} \leq n^k$. ■

Problem 4.5. Prove that the probability that the graph contains a k -clique, for $k \geq 4\log_2 n + 1$, is at most $1/n$.

Solution. To begin with, denote E_S to be the event that S is a k -clique where $|S| = k$. Next, we observe then that the probability of the graph containing a k -clique is simply the union of possible E_S , where S are all of the subsets of size k of V . In other words, we have:

$$\Pr[\text{graph contains a } k\text{-clique}] = \Pr\left[\bigcup E_S\right]$$

And from here, we know that $\Pr[\bigcup E_S] \leq \sum \Pr[E_S] = \frac{\binom{n}{k}}{2^{\binom{k}{2}}}$, which we get from 4.2 and the fact that there are $\binom{n}{k}$ ways of picking a subset S of size k from V .

From here, we observe the following:

$$\begin{aligned} \frac{\binom{n}{k}}{2^{\binom{k}{2}}} &= \frac{\binom{n}{k}}{2^{\frac{k(k-1)}{2}}} \\ &\leq \frac{n^k}{\left(2^{\frac{k-1}{2}}\right)^2} && \text{by inequality in 4.4} \\ &\leq \frac{n^k}{\left(2^{\frac{(4\log_2 n + 1) - 1}{2}}\right)^2} && \text{substitute in } 4\log_2 n + 1 \text{ for } k \\ &= \frac{n^k}{(2^{2\log_2 n})^k} \\ &= \frac{n^k}{n^{2k}} \\ &= \frac{1}{n^k} \\ &\leq \frac{1}{n} \end{aligned}$$

Thus, we see that the probability that the graph contains a k -clique is at most $\frac{1}{n}$. ■

5 PIE Extended

Problem 5.1. One interesting result yielded by the Principle of Inclusion and Exclusion (PIE) is that for any events A_1, A_2, \dots, A_n in some probability space,

$$\sum_{i=1}^n \Pr[A_i] - \sum_{i < j \leq n} \Pr[A_i \cap A_j] + \sum_{i < j < k \leq n} \Pr[A_i \cap A_j \cap A_k] - \dots + (-1)^{n-1} \Pr[A_1 \cap A_2 \cap \dots \cap A_n] \geq 0$$

(Note the LHS is equal to $\Pr[\bigcup_{i=1}^n A_i]$ by PIE, and probability is nonnegative).

Prove that for any events A_1, A_2, \dots, A_n in some probability space,

$$\sum_{i=1}^n \Pr[A_i] - 2 \sum_{i < j \leq n} \Pr[A_i \cap A_j] + 4 \sum_{i < j < k \leq n} \Pr[A_i \cap A_j \cap A_k] - \dots + (-2)^{n-1} \Pr[A_1 \cap A_2 \cap \dots \cap A_n] \geq 0$$

Solution. To begin with, let event B represent an odd number of A_1, \dots, A_n occurring. We observe then that $\Pr[B]$ is:

$$\Pr[B] = \Pr[A_1 \triangle A_2 \triangle \dots \triangle A_n]$$

Proof. We can show that $B = A_1 \triangle \dots \triangle A_n$ is true by induction and showing first that $x \in (A_1 \triangle \dots \triangle A_n) \iff x$ being in an odd number of A_i 's.

Base Case: First, we observe that for $n = 1$, $x \in A_1$; it's in exactly 1 set.

For $n = 2$, we observe that $x \in A_1$ or $x \in A_2$ since $A_1 \triangle A_2 = (A_1 \cup A_2) \setminus (A_1 \cap A_2)$. Thus, it's in an odd number of sets.

Induction Hypothesis: Now, suppose that our claim holds for $n = k$.

Inductive Step: We now want to show that this holds for $n = k + 1$ as well.

Now, suppose $x \in (A_1 \triangle \dots \triangle A_k) \triangle A_{k+1}$. Then, x is in either $A_1 \triangle \dots \triangle A_k$ or x is in A_{k+1} .

If $x \in A_{k+1}$, then we see that it's in exactly one set, and thus our claim is correct.

If it isn't, then from here, we can break it down into two cases.

First, suppose that x is in an even number of sets from A_1 to A_{k+1} . From here, we see that there are two scenarios:

1. If x is in an even number of sets from A_1, \dots, A_k . In this case, x isn't in $A_1 \triangle \dots \triangle A_k$ by our induction hypothesis, and it isn't in A_{k+1} either; thus it isn't in the symmetric difference for A_1, \dots, A_{k+1} .
2. If x is in an odd number of sets from A_1, \dots, A_k . In this case, it must also be in A_{k+1} . But if this is the case, then x can't be inside $A_1 \triangle \dots \triangle A_{k+1}$ by definition.

If x lies in an odd number of sets from A_1 to A_{k+1} then we have the following scenarios:

1. If x lies in an even number of sets from A_1, \dots, A_k and in A_{k+1} as well, we see then that it can't be in the symmetric difference for A_1, \dots, A_{k+1} .
2. If x lies in an odd number of sets from A_1, \dots, A_k , then it can't lie in A_{k+1} . In this case, we see that x lies in the symmetric difference for A_1, \dots, A_{k+1} .

From here, we see then that $x \in (A_1 \triangle \dots \triangle A_k) \triangle A_{k+1}$ if and only if it is in an odd number of the sets. ■

Then, with this in mind, we see then that $\Pr[B] = \Pr[A_1 \triangle \dots \triangle A_n]$.

Now, because $\Pr[B]$ is a probability, it follows then that it must be nonnegative. That is,

$$\Pr[B] = \Pr[A_1 \triangle A_2 \triangle \cdots \triangle A_n] \geq 0.$$

From here, we observe that:

$$\Pr[A_1 \triangle A_2 \triangle \cdots \triangle A_n] = \sum_{i=1}^n \Pr[A_i] - 2 \sum_{i < j \leq n} \Pr[A_i \cap A_j] + \cdots + (-2)^{n-1} \Pr[A_1 \cap A_2 \cap \cdots \cap A_n]$$

We see that this is the case because the RHS is essentially just counting the number of elements in $\Pr[\bigcup_{i=1}^n A_i]$ that are in an odd number of A_i 's. Thus, it follows then that the RHS must be ≥ 0 . ■

6 Independent Complements

Let Ω be a sample space, and let $A, B \subseteq \Omega$ be two independent events.

Problem 6.1. Prove or disprove: \bar{A} and \bar{B} must be independent.

Solution. Suppose that A and B are independent events. Then, by definition, $P(A \cap B) = P(A)P(B)$. From here, observe that $P(\bar{A} \cap \bar{B}) = 1 - (P(A) + P(B) - P(A \cap B))$. From this, we get the following:

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= (1 - P(A))(1 - P(B)) \\ &= P(\bar{A})P(\bar{B}) \end{aligned}$$

And since $P(\bar{A} \cap \bar{B}) = P(\bar{A})P(\bar{B})$, we see that \bar{A} and \bar{B} are independent events as well. ■

Problem 6.2. Prove or disprove: A and \bar{B} must be independent.

Solution. Suppose that A and B are independent events. Then, by definition, $P(A \cap B) = P(A)P(B)$. Now, from here, we observe that $P(A \cap \bar{B}) = P(A) - P(A \cap B)$. With this in mind, we observe the following:

$$\begin{aligned} P(A \cap \bar{B}) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(\bar{B}) \end{aligned}$$

Thus, since $P(A \cap \bar{B}) = P(A)P(\bar{B})$, we see that A and \bar{B} must be independent. ■

Problem 6.3. Prove or disprove: A and \bar{A} must be independent.

Solution. The statement is false.

Suppose that we are flipping a fair coin, with A being the coin landing on heads, and \bar{A} being the coin landing on tails.

We see then that $P(A \cap \bar{A}) = 0$, as it is impossible for a coin to land on both heads and tails simultaneously. However, $P(A)P(\bar{A}) = 0.5(0.5) = 0.25 \neq 0$. Thus, we see that A and \bar{A} aren't independent. ■

Problem 6.4. Prove or disprove: It is possible that $A = B$.

Solution. Suppose that A and B are independent events. Then, we want to show that $P(A \cap A) = P(A)P(A)$. We observe here that $P(A \cap A) = P(A)$. Now, for $P(A) = P(A)P(A)$, the only way this is possible is the following is true:

$$\begin{aligned} P(A) &= P(A)P(A) \\ 0 &= P(A)P(A) - P(A) \\ &= P(A)(P(A) - 1) \end{aligned}$$

In other words, $P(A) = 1$ or $P(A) = 0$. However, because $A, B \subseteq \Omega$, then the only way for this to happen is if A is the sample space itself (which will yield us $P(A) = 1$). ■