CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 0

Due: Friday 8/28, 10:00 PM
Grace period until Friday 8/28 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Administrivia

(a) Make sure you are on the course Piazza (for Q&A) and Gradescope (for submitting homeworks, including this one). Find and familiarize yourself with the course website. What is its homepage's URL?

(b) Read the policies page on the course website.

(i) What is the percentage breakdown of how your grade is calculated (please include both breakdowns)?

(ii) How many discussions do you need to attend to get full credit for discussion attendance?

(iii) Can you attend a section different from the one you signed up for?

(iv) When are the Vitamins due?

2 Course Policies

Go to the course website and read the course policies carefully. Leave a followup in the Homework 0, Question 2 thread on Piazza if you have any questions. Are the following situations violations of course policy? Write "Yes" or "No", and a short explanation for each.

(a) Alice and Bob work on a problem in a study group. They write up a solution together and submit it, noting on their submissions that they wrote up their homework answers together.

(b) Carol goes to a homework party and listens to Dan describe his approach to a problem on the board, taking notes in the process. She writes up her homework submission from her notes, crediting Dan.

(c) Erin comes across a proof that is part of a homework problem while studying course material. She reads it and then, after she has understood it, writes her own solution using the same approach. She submits the homework with a citation to the website.

(d) Frank is having trouble with his homework and asks Grace for help. Grace lets Frank look at her written solution. Frank copies it onto his notebook and uses the copy to write and submit his homework, crediting Grace.

(e) Heidi has completed her homework using LaTeX. Her friend Irene has been working on a homework problem for hours, and asks Heidi for help. Heidi sends Irene her PDF solution, and Irene uses it to write her own solution with a citation to Heidi.

(f) Joe found homework solutions before they were officially released, and every time he got stuck, he looked at the solutions for a hint. He then cited the solutions as part of his submission.

3 Use of Piazza

Piazza is incredibly useful for Q&A in such a large-scale class. We will use Piazza for all important announcements. You should check it frequently. We also highly encourage you to use Piazza to ask questions and answer questions from your fellow students.

(a) Navigate to the "Index" Piazza post, where you can find links to most resources in the course. Write down the Piazza post number for the Note 1 Thread. (When you see @x on Piazza, where x is a positive integer, then x is the post number of the linked post.)

(b) Read the Piazza Etiquette section of the course policies and explain what is wrong with the following hypothetical student question: "Can someone explain the proof of Theorem XYZ to me?" (Assume Theorem XYZ is a complicated concept.)

(c) When are the weekly posts released? Are they required reading?

4 Timezone

Please fill out the discussion time preference form at bit.ly/fa20cs70dispref. What is your magic word?

5 Academic Integrity

Please write or type out the following pledge in print, and sign it.

> I pledge to uphold the university's honor code: to act with honesty, integrity, and respect for others, including their work. By signing, I ensure that all written homework I submit will be in my own words, that I will acknowledge any collaboration or help received, and that I will neither give nor receive help on any examinations.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 1

Due: Friday, 09/04 at 10:00 PM
Grace period until Friday 09/04 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Calculus Review

(a) Compute a closed-form expression for the value of following summation:

$$\sum_{k=1}^{\infty} \frac{9}{2^k}$$

(b) Use summation notion to write an expression equivalent to the following statement:

The sum of the first n consecutive odd integers, starting from 1

(c) Compute the following integral:

$$\int_0^{\infty} \sin(t)e^{-t}dt$$

(d) Find the maximum value of the following function and determine where it occurs:

$$f(x) = -x \cdot \ln x$$

2 Propositional Practice

In parts (a)-(c), convert the English sentences into propositional logic. In parts (d)-(f), convert the propositions into English. In part (f), let $P(a)$ represent the proposition that a is prime.

(a) There is one and only one real solution to the equation $x^2 = 0$.

(b) Between any two distinct rational numbers, there is another rational number.

(c) If the square of an integer is greater than 4, that integer is greater than 2 or it is less than -2.

(d) $(\forall x \in \mathbb{R})\ (x \in \mathbb{C})$

(e) $(\forall x, y \in \mathbb{Z})(x^2 - y^2 \neq 10)$

(f) $(\forall x \in \mathbb{N})\ [\ (x > 1) \implies (\exists a, b \in \mathbb{N})\ ((a + b = 2x) \wedge P(a) \wedge P(b))\ ]$

3 Tautologies and Contradictions

Classify each statement as being one of the following, where P and Q are arbitrary propositions:

- True for all combinations of P and Q (Tautology)
- False for all combinations of P and Q (Contradiction)
- Neither

Justify your answers with a truth table.

(a) $P \implies (Q \wedge P) \vee (\neg Q \wedge P)$

(b) $(P \vee Q) \vee (P \vee \neg Q)$

(c) $P \wedge (P \implies \neg Q) \wedge (Q)$

(d) $(\neg P \implies Q) \implies (\neg Q \implies P)$

(e) $(\neg P \implies \neg Q) \wedge (P \implies \neg Q) \wedge (Q)$

(f) $(\neg(P \wedge Q)) \wedge (P \vee Q)$

4 Prove or Disprove

For each of the following, either prove the statement, or disprove by finding a counterexample.

(a) $(\forall n \in \mathbb{N})$ if n is odd then $n^2 + 4n$ is odd.

(b) $(\forall a, b \in \mathbb{R})$ if $a + b \leq 15$ then $a \leq 11$ or $b \leq 4$.

(c) $(\forall r \in \mathbb{R})$ if r^2 is irrational, then r is irrational.

(d) $(\forall n \in \mathbb{Z}^+)\ 5n^3 > n!$. (Note: $\mathbb{Z}^+$ is the set of positive integers)

5 Twin Primes

(a) Let $p > 3$ be a prime. Prove that p is of the form $3k + 1$ or $3k - 1$ for some integer k.

(b) *Twin primes* are pairs of prime numbers p and q that have a difference of 2. Use part (a) to prove that 5 is the only prime number that takes part in two different twin prime pairs.

6 Social Network

Consider the same setup as Q2 on the vitamin, where there are n people at a party, and every two people are either friends or strangers. Prove or provide a counterexample for the following statements.

(a) For all cases with $n = 5$ people, there exists a group of 3 people that are either all friends or all strangers.

(b) For all cases with $n = 6$ people, there exists a group of 3 people that are either all friends or all strangers.

7 Preserving Set Operations

For a function f, define the image of a set X to be the set $f(X) = \{y \mid y = f(x) \text{ for some } x \in X\}$. Define the inverse image or preimage of a set Y to be the set $f^{-1}(Y) = \{x \mid f(x) \in Y\}$. Prove the following statements, in which A and B are sets. By doing so, you will show that inverse images preserve set operations, but images typically do not.

Hint: For sets X and Y, $X = Y$ if and only if $X \subseteq Y$ and $Y \subseteq X$. To prove that $X \subseteq Y$, it is sufficient to show that $(\forall x)\ ((x \in X) \implies (x \in Y))$.

(a) $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.

(b) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

(c) $f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B)$.

(d) $f(A \cup B) = f(A) \cup f(B)$.

(e) $f(A \cap B) \subseteq f(A) \cap f(B)$, and give an example where equality does not hold.

(f) $f(A \setminus B) \supseteq f(A) \setminus f(B)$, and give an example where equality does not hold.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 2

Due: Friday 09/11 at 10:00 PM
Grace period until Friday 09/11 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Induction

Prove the following using induction:

(a) For all natural numbers $n > 2$, $2^n > 2n + 1$.

(b) For all positive integers n, $1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$.

(c) For all positive natural numbers n, $\frac{5}{4} \cdot 8^n + 3^{3n-1}$ is divisible by 19.

2 Negative pacman returns

Pacman has had a bit of a wild night, and wakes up feeling a bit under the weather. He starts at some location $(i, j) \in \mathbb{N}^2$ in the third quadrant, and is constrained walk on the infinite 2D grid and stay in the third quadrant (say, by walls along the negative x and negative y axes). Every second he does one of the following (if possible):

(i) Walk one step up, to $(i, j+1)$.

(ii) Walk one step right, to $(i+1, j)$.

For example, if he is at $(-5,0)$, his only option is to walk right to $(-4,0)$; if Pacman is instead at $(-3,-2)$, he could walk either to $(-2,-2)$ or $(-3,-1)$.

Prove by induction that no matter how he walks, he will always reach $(0,0)$ in finite time. (*Hint*: Try starting Pacman at a few small points like $(-2,-1)$ and looking all the different paths he could take to reach $(0,0)$. Do you notice a pattern?)

3 Losing Marbles

Two EECS70 GSIs have inexplicably run out of research topics to pursue, papers to read, or homeworks to create, and so they decide to play an incredibly boring game. (This is EECS after all.)

In the game, there is an urn that contains some number of red marbles (R), green marbles (G), and blue marbles (B). There is also an infinite supply of marbles outside the urn.

When it is a player's turn, the player may either:

(i) Remove one red marble from the urn, and add 3 green marbles.

(ii) Remove two green marbles from the urn, and add 7 blue marbles.

(iii) Remove one blue marble from the urn.

These are the only legal moves. The last player that can make a legal move wins. We play optimally, of course – meaning we always play one of the best possible legal moves.

(a) If the urn contains (R, G, B) red, green, and blue marbles initially, then determine the conditions on R, G, B for the first player to win the game. Prove it. In this case, does it matter what strategy the players use?

Hint: Assign each marble a weight, and argue that at every step, the combined weight will go down by exactly 1.

(b) Prove by induction that, if the urn initially contains a finite number of marbles at the start of the game, then the game will end after a finite number of moves.

4 Nothing Can Be Better Than Something

In the stable matching problem, suppose that some jobs and candidates have hard requirements and might not be able to just settle for anything. In other words, in addition to the preference orderings they have, they prefer being unmatched to being matched with some of the lower-ranked entities (in their own preference list). We will use the term entity to refer to a candidate/job. A matching could ultimately have to be partial, i.e., some entities would and should remain unmatched.

Consequently, the notion of stability here should be adjusted a little bit to capture the autonomy of both jobs to unilaterally fire employees and employees to just walk away. A matching is stable if

- there is no matched entity who prefers being unmatched over being with their current partner;
- there is no matched/filled job and unmatched candidate that would both prefer to be matched with each other over their current status;
- similarly, there is no unmatched job and matched candidate that would both prefer to be matched with each other over their current status;

- there is no matched job and matched candidate that would both prefer to be matched with each other over their current partners; and
- there is no unmatched job and unmatched candidate that would both prefer to be with each other over being unmatched.

(a) Prove that a stable pairing still exists in the case where we allow unmatched entities.

(HINT: You can approach this by introducing imaginary/virtual entities that jobs/candidates "match" if they are unmatched. How should you adjust the preference lists of jobs/candidates, including those of the newly introduced imaginary ones for this to work?)

(b) As you saw in the lecture, we may have different stable matchings. But interestingly, if an entity remains unmatched in one stable matching, it/she must remain unmatched in any other stable matching as well. Prove this fact by contradiction.

5 The Ranking List

Let's study the stable matching problem a little bit quantitavely. Here we define the following notation: on day j, let $P_j(M)$ be the rank of the job that applicant M proposes to (where the first application on her list has rank 1 and the last has rank n). Also, let $R_j(W)$ be the total number of applicants that job W has rejected up through day $j-1$ (i.e. not including the proposals on day j). Answer the following questions using the notation above.

(a) Prove or disprove the following claim: $\sum_M P_j(M) - \sum_W R_j(W)$ is independent of j. If it is true, also give the value of $\sum_M P_j(M) - \sum_W R_j(W)$. The notation, $\sum_M$ and $\sum_W$, simply means that we are summing over all applicants and all jobs.

(b) Prove or disprove the following claim: one of the **applicants or jobs** must be matched to something that is ranked in the top half of their preference list. You may assume that n is even.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 3

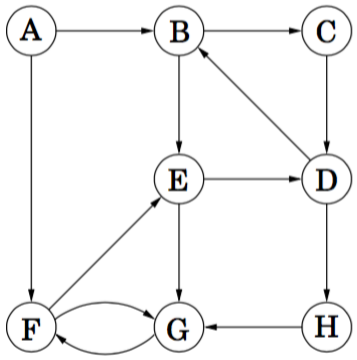
Due: Sunday 09/20 at 10:00 PM
Grace period until Sunday 09/20 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Graph Basics

In the first few parts, you will be answering questions on the following graph G.



(a) What are the vertex and edge sets V and E for graph G?

(b) Which vertex has the highest in-degree? Which vertex has the lowest in-degree? Which vertices have the same in-degree and out-degree?

(c) What are the paths from vertex B to F, assuming no vertex is visited twice? Which one is the shortest path?

(d) Which of the following are cycles in G?

i. $(B,C),(C,D),(D,B)$
ii. $(F,G),(G,F)$
iii. $(A,B),(B,C),(C,D),(D,B)$

iv. $(B,C),(C,D),(D,H),(H,G),(G,F),(F,E),(E,D),(D,B)$

(e) Which of the following are walks in G?

i. (E,G)
ii. $(E,G),(G,F)$
iii. $(F,G),(G,F)$
iv. $(A,B),(B,C),(C,D),(H,G)$
v. $(E,G),(G,F),(F,G),(G,C)$
vi. $(E,D),(D,B),(B,E),(E,D),(D,H),(H,G),(G,F)$

(f) Which of the following are tours in G?

i. (E,G)
ii. $(E,G),(G,F)$
iii. $(F,G),(G,F)$
iv. $(E,D),(D,B),(B,E),(E,D),(D,H),(H,G),(G,F)$

In the following three parts, let's consider a general undirected graph G with n vertices ($n \geq 3$).

(g) True/False: If each vertex of G has degree at most 1, then G does not have a cycle.

(h) True/False: If each vertex of G has degree at least 2, then G has a cycle.

(i) True/False: If each vertex of G has degree at most 2, then G is not connected.

2 Binary Trees

You have seen the recursive definition of binary trees from lecture and from previous classes. Here, we define binary trees in graph theoretic terms as follows (**Note:** here we will modify the definition of leaves slightly for consistency).

- A binary tree of height > 0 is a tree where exactly one vertex, called the **root**, has degree 2, and all other vertices have degrees 1 or 3. Each vertex of degree 1 is called a **leaf**. The **height** h is defined as the maximum length of the path between the root and any leaf.
- A binary tree of height 0 is the graph with a single vertex. The vertex is both a leaf and a root.

(a) Let T be a binary tree of height > 0, and let $h(T)$ denote it's height. Let r be the root in T and u and v be it's neighbors. Show that removing r from T will result in two binary trees, L,R with roots u and v respectively. Also, show that $h(T) = \max(h(L),h(R)) + 1$

(b) Using the graph theoretic definition of binary trees, prove that the number of vertices in a binary tree of height h is at most $2^{h+1} - 1$

(c) Prove that all binary trees with n leaves have $2n - 1$ vertices

3 Proofs in Graphs

Please prove or disprove the following claims.

(a) On the axis from San Francisco traffic habits to Los Angeles traffic habits, Old California is more towards San Francisco: that is, civilized. In Old California, all roads were one way streets. Suppose Old California had n cities ($n \geq 2$) such that for every pair of cities X and Y, either X had a road to Y or Y had a road to X. Prove or disprove that there existed a city which was reachable from every other city by traveling through at most 2 roads.

[*Hint:* Induction]

(b) In lecture, we have shown that a connected undirected graph has an Eulerian tour if and only if every vertex has even degree.

Consider a connected graph G with n vertices which has exactly $2m$ vertices of odd degree, where $m > 0$. Prove or disprove that there are m walks that *together* cover all the edges of G (i.e., each edge of G occurs in exactly one of the m walks, and each of the walks should not contain any particular edge more than once).

4 Planarity

(a) Prove that $K_{3,3}$ is nonplanar.

(b) Consider graphs with the property T: For every three distinct vertices v_1, v_2, v_3 of graph G, there are at least two edges among them. Use a proof by contradiction to show that if G is a graph on ≥ 7 vertices, and G has property T, then G is nonplanar.

5 Always, Sometimes, or Never

In each part below, you are given some information about the so-called original graph, OG. Using only the information in the current part, say whether OG will always be planar, always be non-planar, or could be either. If you think it is always planar or always non-planar, prove it. If you think it could be either, give a planar example and a non-planar example.

(a) OG can be vertex-colored with 4 colors.

(b) OG requires 7 colors to be vertex-colored.

(c) $e \leq 3v - 6$, where e is the number of edges of OG and v is the number of vertices of OG.

(d) OG is connected, and each vertex in OG has degree at most 2.

(e) Each vertex in OG has degree at most 2.

6 Touring Hypercube

In the lecture, you have seen that if G is a hypercube of dimension n, then

- The vertices of G are the binary strings of length n.
- u and v are connected by an edge if they differ in exactly one bit location.

A *Hamiltonian tour* of a graph is a sequence of vertices $v_0, v_1, \ldots, v_k$ such that:

- Each vertex appears exactly once in the sequence.
- Each pair of consecutive vertices is connected by an edge.
- v_0 and v_k are connected by an edge.

(a) Show that a hypercube has an Eulerian tour if and only if n is even. (*Hint: Euler's theorem*)

(b) Show that every hypercube has a Hamiltonian tour.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 4

Due: Friday 09/25 at 10:00 PM
Grace period until Friday 09/25 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Fibonacci GCD

The Fibonacci sequence is given by $F_n = F_{n-1} + F_{n-2}$, where $F_0 = 0$ and $F_1 = 1$. Prove that, for all $n \geq 0$, $\gcd(F_n, F_{n-1}) = 1$.

2 The Last Digit

In each case show your work and justify your answers.

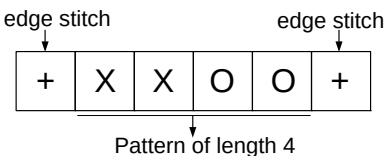
(a) If $9k+5$ and $2k+1$ have the same last digit for some natural number k, find the last digit of k.

(b) If $S = \sum_{i=1}^{19} i!$, then find the last digit of S^2.

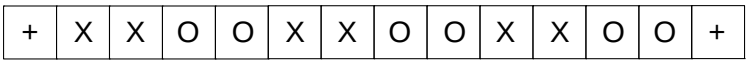
3 Celebrate and Remember Textiles

Mathematics and computing both owe an immense debt to textiles, where many key ideas originated.

Instructions for knitting patterns will tell you to begin by "casting on" the needle some multiple of m plus r, where m is the number of stitches to create one repetition of the pattern and r is the number of stitches needed for the two edges of the piece. For example, in the simple rib stitch pattern below, the repeating pattern is of length $m = 4$, and you need $r = 2$ stitches for the edges.



Thus, to make the final piece wider, you can add as many multiples of the pattern of length 4 as you like; for example, if you want to repeat the pattern 3 times, you need to cast on a total of $3m + r = 3(4) + 2 = 14$ stitches (shown below).



You've decided to knit a 70-themed baby blanket as a gift for your cousin and want to incorporate rows from three different stitch patterns with the following requirements:

- Alternating Link: Multiple of 7, plus 4
- Double Broken Rib: Multiple of 4, plus 2
- Swag: Multiple of 5, plus 2

You want to be able to switch between knitting these different patterns without changing the number of stitches on the needle, so you must use a number of stitches that simultaneously meets the requirements of all three patterns. **Find the smallest number of stitches you need to cast on in order to incorporate all three patterns in your baby blanket.**

4 Sparsity of Primes

A prime power is a number that can be written as p^i for some prime p and some positive integer i. So, $9 = 3^2$ is a prime power, and so is $8 = 2^3$. $42 = 2 \cdot 3 \cdot 7$ is not a prime power.

Prove that for any positive integer k, there exists k consecutive positive integers such that none of them are prime powers.

Hint: this is a Chinese Remainder Theorem problem

5 Fermat's Little Theorem

Fermat's Little Theorem states that for any prime p and any $a \in \{1, 2, \ldots, p-1\}$, we have $a^{p-1} \equiv 1 \pmod{p}$. Without using induction, prove that $\forall n \in \mathbb{N}$, $n^7 - n$ is divisible by 42.

6 A Taste of RSA

Suppose that p and q are distinct odd primes (i.e. they are primes > 2). Define $N = pq$. Let a be any integer that is relatively prime to N. In other words, $\gcd(a, N) = 1$. Prove that $a^{(p-1)(q-1)+1} \equiv a \pmod{N}$. It turns out that this equivalence is in fact the basis of RSA, as you will see soon in class.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 5

Due: Friday 10/02 at 10:00 PM
Grace period until Friday 10/02 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 RSA with Just One Prime

Given the message $x \in \{0, 1, \ldots, N-1\}$ and $N = pq$, where p and q are prime numbers, conventional RSA encrypts x with $y = E(x) \equiv x^e \pmod{N}$. The decryption is done by $D(y) \equiv y^d \pmod{N}$, where d is the inverse of $e \pmod{(p-1)(q-1)}$.

Alice is trying to send a message to Bob, and as usual, Eve is trying to decipher what the message is. One day, Bob gets lazy and tells Alice that he will now use $N = p$, where p is a 1024-bit prime number, as part of his public key. He tells Alice that it's okay, since Eve will have to try out 2^{1024} combinations to guess x. It is very likely that Eve will not find out the secret message in a reasonable amount of time! In this problem, we will see whether Bob is right or wrong. Assume that Eve has found out about this new setup and that she knows the public key.

Similar to the original method, for any message $x \in \{0, 1, \ldots, N-1\}$, $E(x) \equiv x^e \pmod{p}$, and $D(y) \equiv y^d \pmod{p}$. Choose e such that it is coprime with $p-1$, and choose $d \equiv e^{-1} \pmod{p-1}$.

(a) Prove that the message x is recovered after it goes through your new encryption and decryption functions, $E(x)$ and $D(y)$.

(b) Can Eve compute d in the decryption function? If so, by what algorithm and approximately how many iterations does it take for it to terminate?

(c) Given part (b), how would Eve recover x and what algorithm would she use? Approximately how many iterations does it take to terminate?

(d) Based on the previous parts, can Eve recover the original message in a reasonable amount of time? Explain.

2 Squared RSA

(a) Prove the identity $a^{p(p-1)} \equiv 1 \pmod{p^2}$, where a is coprime to p, and p is prime. (Hint: Try to mimic the proof of Fermat's Little Theorem from the notes.)

(b) Now consider the RSA scheme: the public key is $(N = p^2q^2, e)$ for primes p and q, with e relatively prime to $p(p-1)q(q-1)$. The private key is $d = e^{-1} \pmod{p(p-1)q(q-1)}$. Prove that the scheme is correct for x relatively prime to both p and q, i.e. $x^{ed} \equiv x \pmod{N}$. (Hint: Try to mimic the proof of RSA correctness from the notes.)

3 The CRT and Lagrange Interpolation

Let $n_1, \ldots n_k$ be pairwise co-prime, i.e. n_i and n_j are co-prime for all $i \neq j$. The Chinese Remainder Theorem (CRT) tells us that there exist solutions to the following system of congruences:

$$x \equiv a_1 \pmod{n_1} \tag{1}$$
$$x \equiv a_2 \pmod{n_2} \tag{2}$$
$$\vdots \tag{$\vdots$}$$
$$x \equiv a_k \pmod{n_k} \tag{k}$$

and all solutions are equivalent $\pmod{n_1 n_2 \cdots n_k}$. For this problem, parts (a)-(c) will walk us through a proof of the Chinese Remainder Theorem. We will then use the CRT to revisit Lagrange interpolation.

(a) We start by proving the $k = 2$ case: Prove that we can always find an integer x_1 that solves (1) and (2) with $a_1 = 1, a_2 = 0$. Similarly, prove that we can always find an integer x_2 that solves (1) and (2) with $a_1 = 0, a_2 = 1$.

(b) Use part (a) to prove that we can always find at least one solution to (1) and (2) for any a_1, a_2. Furthermore, prove that all possible solutions are equivalent $\pmod{n_1 n_2}$.

(c) Now we can tackle the case of arbitrary k: Use part (b) to prove that there exists a solution x to (1)-(k) and that this solution is unique $\pmod{n_1 n_2 \cdots n_k}$.

For polynomials $p_1(x)$, $p_2(x)$ and $q(x)$ we say that $p_1(x) \equiv p_2(x) \bmod q(x)$ if $p_1(x) - p_2(x)$ is of the form $q(x) \times m(x)$ for some polynomial $m(x)$.

(d) Define the polynomials $x - a$ and $x - b$ to be co-prime if they have no common divisor of degree 1. Assuming that the CRT still holds when replacing x, a_i and n_i with polynomials (using the definition of co-prime polynomials just given), show that the system of congruences

$$p(x) \equiv y_1 \pmod{(x - x_1)} \tag{1'}$$
$$p(x) \equiv y_2 \pmod{(x - x_2)} \tag{2'}$$
$$\vdots \tag{$\vdots$}$$
$$p(x) \equiv y_k \pmod{(x - x_k)} \tag{k'}$$

has a unique solution $\pmod{(x-x_1)\cdots(x-x_k)}$ whenever the x_i are pairwise distinct. What is the connection to Lagrange interpolation?

Hint: To show that a unique solution exists, you may use the fact that the CRT has a unique solution when certain properties are satisfied.

4 Polynomials in Fields

Define the sequence of polynomials by $P_0(x) = x + 12$, $P_1(x) = x^2 - 5x + 5$ and $P_n(x) = xP_{n-2}(x) - P_{n-1}(x)$.

(For instance, $P_2(x) = 17x - 5$ and $P_3(x) = x^3 - 5x^2 - 12x + 5$.)

(a) Show that $P_n(7) \equiv 0 \pmod{19}$ for every $n \in \mathbb{N}$.

(b) Show that, for every prime q, if $P_{2017}(x) \not\equiv 0 \pmod{q}$, then $P_{2017}(x)$ has at most 2017 roots modulo q.

5 Secrets in the United Nations

A vault in the United Nations can be opened with a secret combination $s \in \mathbb{Z}$. In only two situations should this vault be opened: (i) all 193 member countries must agree, or (ii) at least 55 countries, plus the U.N. Secretary-General, must agree.

(a) Propose a scheme that gives private information to the Secretary-General and all 193 member countries so that the secret combination s can only be recovered under either one of the two specified conditions.

(b) The General Assembly of the UN decides to add an extra level of security: each of the 193 member countries has a delegation of 12 representatives, all of whom must agree in order for that country to help open the vault. Propose a scheme that adds this new feature. The scheme should give private information to the Secretary-General and to each representative of each country.

6 Secret Sharing with Spies

An officer stored an important letter in her safe. In case she is killed in battle, she decides to share the password (which is a number) with her troops. However, everyone knows that there are 3 spies among the troops, but no one knows who they are except for the three spies themselves. The 3 spies can coordinate with each other and they will either lie and make people not able to open the safe, or will open the safe themselves if they can. Therefore, the officer would like a scheme to share the password that satisfies the following conditions:

- When M of them get together, they are guaranteed to be able to open the safe even if they have spies among them.

- The 3 spies must not be able to open the safe all by themselves.

Please help the officer to design a scheme to share her password. What is the scheme? What is the smallest M? Show your work and argue why your scheme works and any smaller M couldn't work. (The troops only have one chance to open the safe; if they fail the safe will self-destruct.)

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 06

Due: Friday 10/09 at 10:00 PM
Grace period until Friday 10/09 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Exam Policy and Practice

Please read the all the documents (Policy, Key Changes, and Reminders) of the Exam Policy carefully before proceeding. This question is designed to familiarize you with some of the things you will have to do during the exam.

(a) After reading through the Exam Policy carefully, please answer the following questions.

(i) Given you experience no disruptions during the exam, how many total minutes do you have for scanning and submission?

(ii) Are you required to record locally during the exam? How much space should you have available on your computer for a local recording?

(iii) How should you contact the course staff for an emergency situation during the exam?

(b) Please configure your Zoom link.

(i) You should use the same Zoom link to join the meeting for the midterm as the Zoom link that you send to us. This can easily be done by submitting your Personal Meeting Room link and setting your Personal Meeting ID as your default on all devices you will be using for the final.

(ii) Ensure anyone can join your Zoom link and that there is no waiting room for your Zoom meeting.

(iii) Please the following Google Form with your Zoom link that you plan to use.

(c) You will now conduct a Zoom recording. Please read all instructions beforehand. You will use this recording to submit the mock midterm on gradescope, and should use the remaining time of the recording to work through a practice exam or other study material to simulate the actual circumstances of the final exam. It is advised to complete the LaTex Rehearsal beforehand, to familiarize yourself with typing LaTex answers.

(i) Start the Zoom call for the link you provided above. Turn on your microphone and recording device (webcam, phone camera). Turn off your speaker. Share your entire desktop (not just a particular window).

(ii) Start recording via Zoom. You may record locally or on the cloud.

(iii) Hold your CalID next to your face and record yourself saying your name into the webcam. Both your face and your entire CalID should be visible in the video. We should be able to read your name and SID. This step should take **at least** 3 seconds. See figure **??**. *If you do not have a CalID for some reason, please hold up some document which has an image of you and proves your identity, such as a driver's license.*

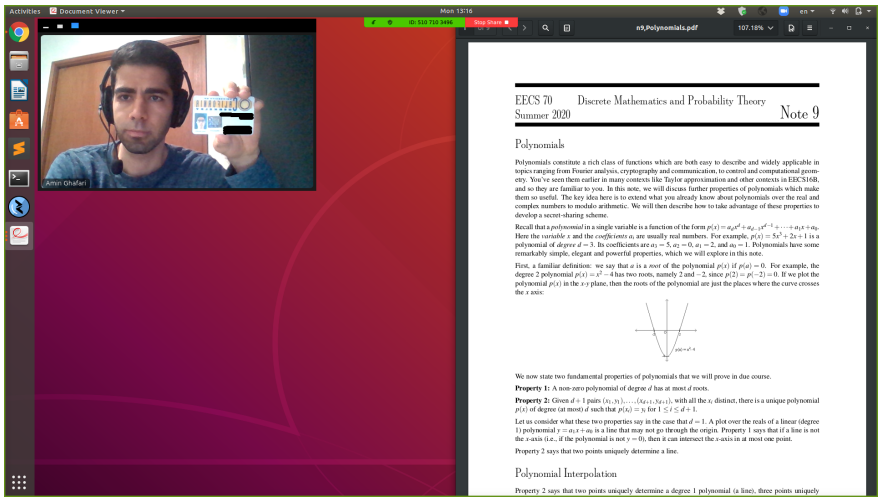


Figure 1: ID card demonstration. Do not actually black out your SID and name.

(iv) Position your recording device in such a way that we can see your workspace and your hands as best as possible. We suggest using your phone to record your hands, but if you are not, then it should be visible in the recording, face down. See figure **??**.

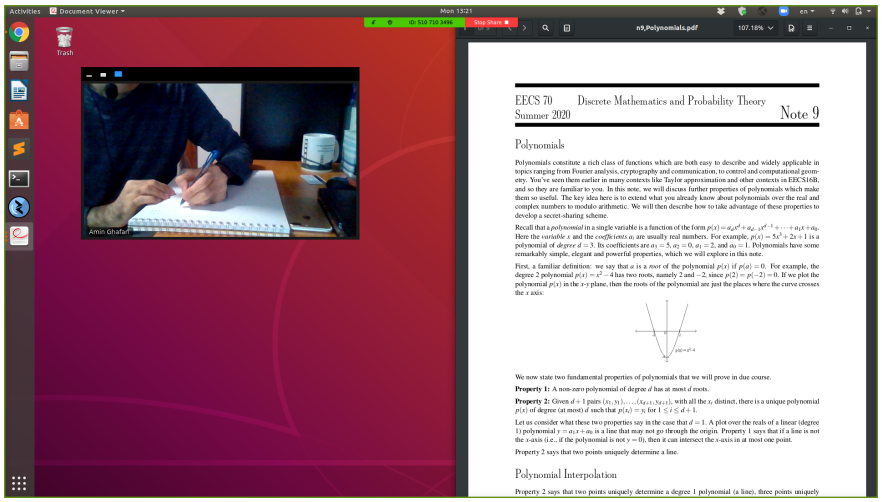


Figure 2: Demonstration of taking your exam. Your setup should look like this while you are taking the exam. The video must be on with your hands visible, alongside your exam pdf or gradescope.

(v) Your microphone should be on at all times. We should be able to see the time on your desktop at all times.

(vi) Record for two hours.

(vii) There are two mock midterm assignments on gradescope. You will see similar assignments on the day of the actual midterm. The one with (Short Answer) will be similar to Vitamins, where you will enter your answers in the gradescope assignment. The other (Written), will be full solution questions, and you will need to scan in your answers.

(viii) Complete and submit to the two mock midterm assignments. This includes both the short answers online, as well as scanning and submitting the written portion of the assignment.

(ix) For the remaining time, you should work through a practice final exam or other study material for the course. The more realistic it is to actually taking a final, the better practice it will be for you on the final.

(x) After two hours, stop the recording. Check your recording to confirm that it contains your video of your hands as well as your desktop throughout its duration. Upload your video to Google drive and submit the link to the video using this Google Form. You must make sure that the link sharing permissions are set so that we may view the video. Write down the magic words from the Google Form. DO NOT use this form for the actual exam, refer to the link in the policies during the actual midterm.

Link for policy:
https://docs.google.com/document/d/1-r3KrjQ46lX-OOiwx6lsoqAeSW0bDdM2oIw0xlKhLi0/edit?usp=sharing

Form to submit Zoom link:
https://forms.gle/2HDJtQijTzQdutgX8

Form to submit 2 hour video link:
https://forms.gle/XTJLhhbhqBNnKkxN9

2 Message is too noisy

In this problem, we are going to discuss the decoding procedure even when the codeword is corrupted more than they could be. For all parts, work in mod 17.

(a) Encode the message $(0,1,4)$ into a polynomial, where $P(0) = 0, P(1) = 1, P(2) = 4$, what is P?

(b) Suppose you send the message $(P(0), P(1), P(2), P(3), P(4))$ to the receiver and last packet is corrupted to 0. Run the decoding process and calulate the Q, E as defined in the lecture. You should also confirm that $Q(x)/E(x) = P(x)$.

(c) After corrupting the 4-th packet to 6 and 5-th packet to 8, decode again, by computing $Q, E, Q(x)/E(x)$, and outputting the first 3 packets. Explain why the decoded message is not the original message, but rather $(1,1,4)$.

(d) Define the Hamming distance between two messages to be the number of packets that differ. For example, the distance between $(0,1,2,3,4)$ and $(0,1,1,4,4)$ is 2 since they differ at the third and forth position.

Let $RS[5,3]$ be all Reed-Solomon codewords with codeword length 5, message length 3. Show that the Hamming distance between any two codewords in $RS[5,3]$ is at least 3. Also show that the codeword $(1,1,3,7,13)$ (which the decoder finds) has the smallest Hamming distance from the non-codeword $(1,1,4,7,13)$ compared to all other codeword in $RS[5,3]$.

(e) We generalize $RS[m,n]$ to be all Reed-Solomon codewords with length m, message length n. (Note: min Hamming distance between any pair of valid codewords is $m-n+1$). Let C' be the corrupted codeword, $msg = Decode(C')$, $E = Encode(msg)$. $Hamming(x,y)$ is the hamming distance between x and y. Show

$$Hamming(C',E) = \min_{E' \in RS[m,n]} (Hamming(C',E'))$$

Hint: if there are too many corruptions, clearly it will decode to a wrong message.

3 Linearity

Prove that Reed Solomon codes are *linear*; that is, the element-wise sum of two Reed Solomon codewords is also a Reed Solomon codeword. To do this, use the coefficient encoding rather than interpolation encoding: If you have a message of length n and you want to send m packets, create a degree $n-1$ polynomial $p(x)$ where your message $(c_0, c_1, \ldots, c_{n-1})$ are the coefficients of $p(x)$, and the codeword is the evaluation of $p(x)$ at $\{0,1,\ldots,m-1\}$. (Assume we are working on $GF(p)$ for large enough p.)

4 Multiplicative

Recall $RS[m,n]$ to be all Reed-Solomon codewords with length m, message length n. Given two codewords $a,b \in RS[m,n]$. Let $c = a * b$ be the element-wise product of a and b. Show that $c \in RS[m,2n-1]$. (Assume we are working over $GF(p)$, where p is large enough)

5 Maze

(a) Given a 4×4 grid, how many different paths from $(0,0)$ to $(4,4)$ satisfy the following condition:

- You can only go from (x,y) to either $(x+1,y)$ or $(x,y+1)$

(b) Given a 4×4 grid, how many different paths from $(0,0)$ to $(4,4)$ satisfy the following condition:

- You can only go from (x,y) to either $(x+1,y)$ or $(x,y+1)$

- You cannot go to points (x, y) where $y > x$, in other word, you cannot cross line $y = x$

(c) How many sequences of 4 pairs of parentheses are mismatched? An example of a matched sequence of paretheses is ()()()(), while a mismatched sequence is)))(((.

6 Good Game

Player will send 'GG' (Good Game) to the winner after each defeat in a 1v1 competitive game. Maru is a skilled Terran player in the Game. And he is 27 pts behind Player Sierral. Suppose Sierral has finished all of his games and Maru has 10 games to go. If Maru wins the i-th game, he will get i pts.

(a) What's the maximum number of GGs that Maru can send and have a higher points than Sierral?

(b) How many different ways that Maru can defeat Sierral (earns more than 27 pts)?

7 Counting Functions

(a) Compute $g(n)$, the number of ways to divide $\{1, 2, 3, ..., n\}$ into 2 non-empty groups.

(b) Compute $f(n)$, the number of ways to divide $\{1, 2, 3, ..., n\}$ into 3 non-empty groups. (Hint: our calculation involves a recursive formula, and included g)

(c) How many surjective functions $h : \{1, 2, 3, ..., 7\} \to \{1, 2, 3\}$? You may leave your answer in terms of f and g for partial credit, but also compute the actual number.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 7

Due: Friday 10/16 at 10:00 PM
Grace period until Friday 10/16 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Strings

How many different strings only contains A, B, C? And how many such strings contains at least one of each characters?

2 Palindromes

How many 5-digit palindromes are there? (A palindrome is a number that reads the same way forwards and backwards. For example, 27872 and 48484 are palindromes, but 28389 and 12541 are not.)

3 Maze in general and Trees too!!!

Given an maze of sidelength n where one starts at $(0,0)$ and goes to (n,n).

(a) How many shortest paths are there that go from $(0,0)$ to (n,n)?

(b) Extending the width by 1, how many shortest paths are there that go from $(0,0)$ to $(n-1,n+1)$.

(c) Now consider shortest paths that meet the conditions which only use to points (x,y) where $y \leq x$. That is, the path cannot cross line $y = x$.

i. Give an expression using part (a) and (b), that counts the number of paths. (Hint: consider what happens after a shortest that crosses $y = x$ at (i,i), that is, the remaining path starting from $(i,i+1)$ and then continuing to (n,n). If in the remainder of the path, one exchanges the y-direction moves with x-direction moves and vice versa, where does one end up?

ii. A different tack is to derive a recursive formula. We call these paths n-legal paths for a maze of sidelength n, and let F_n be the number of n-legal paths.

Consider a path, and let $i < n$ be the largest value where the path contains (i,i), argue the number of paths is then $F_i * F_{n-i-1}$.

(Hint: if $i = 0$, what are your first and last moves, and where is the remainder of the path allowed to go.)

iii. Give a recursive formula for the number of spanning trees of a complete graph K_n for $n \geq 3$, where each non-root node has degree 3 or 1, and at most 1 node has degree 2?

Two trees are different if and only if either left-subtree is different or right-subtree is different.

(Notice something about your formula and the maze problem. Neat!)

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 8

Due: Friday 10/23 at 10:00 PM
Grace period until Friday 10/23 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Countability Proof Practice

(a) A disk is a 2D region of the form $\{(x,y) \in \mathbb{R}^2 : (x-x_0)^2 + (y-y_0)^2 \leq r^2\}$, for some $x_0, y_0, r \in \mathbb{R}$, $r > 0$. Say you have a set of disks in $\mathbb{R}^2$ such that none of the disks overlap. Is this set always countable, or potentially uncountable?
(*Hint*: Attempt to relate it to a set that we know is countable, such as $\mathbb{Q} \times \mathbb{Q}$)

(b) A circle is a subset of the plane of the form $\{(x,y) \in \mathbb{R}^2 : (x-x_0)^2 + (y-y_0)^2 = r^2\}$ for some $x_0, y_0, r \in \mathbb{R}$, $r > 0$. Now say you have a set of circles in $\mathbb{R}^2$ such that none of the circles overlap. Is this set always countable, or potentially uncountable?
(*Hint*: The difference between a circle and a disk is that a disk contains all of the points in its interior, whereas a circle does not.)

(c) Is the set containing all increasing functions $f : \mathbb{N} \to \mathbb{N}$ (i.e., if $x \geq y$, then $f(x) \geq f(y)$) countable or uncountable? Prove your answer.

(d) Is the set containing all decreasing functions $f : \mathbb{N} \to \mathbb{N}$ (i.e., if $x \geq y$, then $f(x) \leq f(y)$) countable or uncountable? Prove your answer.

2 Hilbert's Hotel

You don't have any summer plans, so you decide to spend a few months working for a magical hotel with a countably infinite number of rooms. The rooms are numbered according to the natural numbers, and all the rooms are currently occupied. Assume that guests don't mind being moved from their current room to a new one, so long as they can get to the new room in a finite amount of time (i.e. guests can't be moved into a room infinitely far from their current one).

(a) A new guest arrives at the hotel. All the current rooms are full, but your manager has told you never to turn away a guest. How could you accommodate the new guest by shuffling other guests around? What if you instead had k guest arrive, for some fixed, positive $k \in \mathbb{Z}$?

(b) Unfortunately, just after you've figured out how to accommodate your first $k+1$ guests, a countably infinite number of guests arrives in town on an infinitely long train. The guests on the train are sitting in seats numbered according to the natural numbers. How could you accommodate all the new guests?

(c) Thanks to a (literally) endless stream of positive TripAdvisor reviews, word of the infinite hotel gets around quickly. Soon enough you find out that a countably infinite number of trains have arrived in town. Each is of infinite length, and carries a countably infinite number of passengers. How would you accommodate all the new passengers?

3 Finite and Infinite Graphs

The graph material that we learned in lecture still applies if the set of vertices of a graph is infinite. We thus make a distinction between finite and infinite graphs: a graph $G = (V, E)$ is finite if V and E are both finite. Otherwise, the graph is infnite. As examples, consider the graphs

- $G_1 = (V = \mathbb{Z},\ E = \{(i, j) \in \mathbb{Z} \times \mathbb{Z} |\ |i - j| = 1\})$
- $G_2 = (V = \mathbb{Z},\ E = \{(i, j) \in \mathbb{Z} \times \mathbb{Z} |\ i < j\})$
- $G_3 = (V = \mathbb{Z}^2,\ E = \{((i, j), (k, l)) \in \mathbb{Z}^2 \times \mathbb{Z}^2 |\ (i = k \wedge |j - l| = 1) \vee (j = l \wedge |i - k| = 1)\})$

Observe that G_1 is a line of integers, G_2 is a complete graph over all integers, and G_3 is an grid of integers. Prove whether the following sets of graphs are countable or uncountable

(a) The set of all finite graphs $G = (V, E)$, for $V \subseteq \mathbb{N}$

(b) The set of all infinite graphs over a fixed, countably infinite set of vertices (in other words, they all have the same vertex set).

(c) The set of all graphs over a fixed, countably infinite set of vertices, the degree of each vertex is exactly two. For instance, every vertex in G_1 (defined above) has degree 2.
(*Hint:* if V is a countably infinite set, then the set of bijections $f : V \to V$ is uncountable. You may use this fact without proof.)

(d) We say that graphs $G = (V, E)$ and $G' = (V', E')$ are isomorphic if the exists some bijection $f : V \to V'$ such that $(u, v) \in V$ iff $(f(u), f(v)) \in V'$. Such a bijection f is called a **graph isomorphism**. Suppose we consider two graphs to be the equivalent if they are isomorphic. The idea is that if we relabel the vertices of a graph, it is still the same graph. Using this definition of "being the same graph", can you conclude that the set of trees over countably infinite vertices is countable?
(*Hint:* Begin by showing that for any graph isomorphism f, and any vertex v, $f(v)$ and v have the same degree)

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 9

Due: Friday 10/30 at 10:00 PM
Grace period until Friday 10/30 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Unprogrammable Programs

Prove whether the programs described below can exist or not.

(a) A program $P(F,x,y)$ that returns true if the program F outputs y when given x as input (i.e. $F(x) = y$) and false otherwise.

(b) A program P that takes two programs F and G as arguments, and returns true if F and G halt on the same set of inputs (or false otherwise).

2 Computations on Programs

(a) Is it possible to write a program that takes a natural number n as input, and finds the shortest arithmetic formula which computes n? For the purpose of this question, a formula is a sequence consisting of some valid combination of (decimal) digits, standard binary operators ($+$, $\times$, the "^" operator that raises to a power), and parentheses. We define the length of a formula as the number of characters in the formula. Specifically, each operator, decimal digit, or parentheses counts as one character.

(*Hint:* Think about whether it's possible to enumerate the set of possible arithmetic formulas. How would you know when to stop?)

(b) Now say you wish to write a program that, given a natural number input n, finds another program (e.g. in Java or C) which prints out n. The discovered program should have the minimum execution-time-plus-length of all the programs that print n. Execution time is measured by the number of CPU instructions executed, while "length" is the number of characters in the source code. Can this be done?

(*Hint:* Is it possible to tell whether a program halts on a given input within t steps? What can you say about the execution-time-plus-length of the program if you know that it does not halt within t steps?)

3 Kolmogorov Complexity

Compressing a bit string x of length n can be interpreted as the task of creating a program of fewer than n bits that returns x. The Kolmogorov complexity of a string $K(x)$ is the length of an optimally-compressed copy of x; that is, $K(x)$ is the length of shortest program that returns x.

(a) Explain why the notion of the "smallest positive integer that cannot be defined in under 280 characters" is paradoxical.

(b) Prove that for any length n, there is at least one string of bits that cannot be compressed to less than n bits.

(c) Say you have a program K that outputs the Kolmogorov complexity of any input string. Under the assumption that you can use such a program K as a subroutine, design another program P that takes an integer n as input, and outputs the length-n binary string with the highest Kolmogorov complexity. If there is more than one string with the highest complexity, output the one that comes first alphabetically.

(d) Let's say you compile the program P you just wrote and get an m bit executable, for some $m \in \mathbb{N}$ (i.e. the program P can be represented in m bits). Prove that the program P (and consequently the program K) cannot exist.

(*Hint*: Consider what happens when P is given a very large input n.)

4 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 ($p = 0.5$).

(a) What is the size of the sample space, $|\Omega|$?

(b) How many elements of Ω have exactly three heads?

(c) How many elements of Ω have three or more heads?
(*Hint:* Argue by symmetry.)

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).

(d) What is the probability that you will observe the sequence HHHTT? What about HHHHT?

(e) What is the chance of observing at least one head?

(f) What about the chance of observing three or more heads?

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

(g) What is the chance of observing the outcome HHHTT? What about HHHHT?

(h) What about the chance of at least one head?

(i) What about the chance of ≥ 3 heads?

5 Ball-and-Bin Counting Problems

Say you have 5 bins, and randomly throw 7 balls into them.

1. What is the probability that the first bin has precisely 3 balls in it?
2. What is the probability that the third bin has at least 3 balls in it?
3. What is the probability that at least one of the bins has precisely 3 balls in it?

6 Monty Hall's Revenge

Due to a quirk of the television studio's recruitment process, Monty Hall has ended up drawing all the contestants for his game show from among the ranks of former CS70 students. Unfortunately for Monty, the former students' amazing probability skills have made his cars-and-goats gimmick unprofitable for the studio. Monty decides to up the stakes by asking his contestants to generalise to three new situations with a variable number of doors, goats, and cars:

(a) There are n doors for some $n > 2$. One has a car behind it, and the remaining $n-1$ have goats. As in the ordinary Monty Hall problem, Monty will reveal one door with a goat behind it after you make your first selection. How would switching affect the odds that you select the car?

(Hint: Think about the size of the sample space for the experiment where you *always* switch. How many of those outcomes are favorable?)

(b) Again there are $n > 2$ doors, one with a car and $n-1$ with goats, but this time Monty will reveal $n-2$ doors with goats behind them instead of just one. How does switching affect the odds of winning in this modified scenario?

(c) Finally, imagine there are $k < n-1$ cars and $n-k$ goats behind the $n > 2$ doors. After you make your first pick, Monty will reveal $j < n-k$ doors with goats. What values of j,k maximize the relative improvement in your odds of winning if you choose to switch? (i.e. what j,k maximizes the ratio between your odds of winning when you switch, and your odds of winning when you do not switch?)

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 10

Due: Friday 11/06 at 10:00 PM
Grace period until Friday 11/06 at 11:59PM
Maximum credit for this homework will be given for any score of 50% or more.

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Independent Complements

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

(a) Prove or disprove: $\overline{A}$ and $\overline{B}$ must be independent.

(b) Prove or disprove: A and $\overline{B}$ must be independent.

(c) Prove or disprove: A and $\overline{A}$ must be independent.

(d) Prove or disprove: It is possible that $A = B$.

2 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?

3 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.

(a) What is the probability that the game ends by flipping exactly five coins?

(b) Given that the game ends after flipping the fifth coin, what is the probability that three heads appear in the sequence?

(c) 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?

4 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.)

5 Identity Theft

A group of n friends go to the gym together, and while they are playing basketball, they leave their bags against the nearby wall. An evildoer comes, takes the student ID cards from the bags, randomly rearranges them, and places them back in the bags, one ID card per bag.

(a) What is the probability that no one receives his or her own ID card back?

Hint: Use the inclusion-exclusion principle.

(b) What is the limit of this proability as $n \to \infty$?

Hint: $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

6 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.

(a) What is the probability that exactly k balls land in the first bin, where k is an integer $0 \leq k \leq n$?

(b) 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.)

(c) 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.

(d) 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?

(e) 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.)

7 Cliques in Random Graphs

In last week's homework you worked on a 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:

(a) What is the size of the sample space?

(b) 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?

(c) For two sets of vertices $V_1 = \{v_1, \ldots, v_\ell\}$ and $V_2 = \{w_1, \ldots, w_k\}$, are E_{V_1} and E_{V_2} independent?

(d) Prove that $\binom{n}{k} \leq n^k$.

(e) 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).

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 11

Due: Friday 11/14 at 10:00 PM
Grace period until Friday 11/14 at 11:59PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Random Variables Warm-Up

Let X and Y be random variables, each taking values in the set $\{0,1,2\}$, with joint distribution

$$\begin{array}{lll} \mathbb{P}[X=0,Y=0]=1/3 & \mathbb{P}[X=0,Y=1]=0 & \mathbb{P}[X=0,Y=2]=1/3 \\ \mathbb{P}[X=1,Y=0]=0 & \mathbb{P}[X=1,Y=1]=1/9 & \mathbb{P}[X=1,Y=2]=0 \\ \mathbb{P}[X=2,Y=0]=1/9 & \mathbb{P}[X=2,Y=1]=1/9 & \mathbb{P}[X=2,Y=2]=0. \end{array}$$

(a) What are the marginal distributions of X and Y?

(b) What are $\mathbb{E}[X]$ and $\mathbb{E}[Y]$?

(c) (optional) What are $\mathrm{Var}(X)$ and $\mathrm{Var}(Y)$?

(d) Let I be the indicator that $X=1$, and J be the indicator that $Y=1$. What are $\mathbb{E}[I]$, $\mathbb{E}[J]$ and $\mathbb{E}[IJ]$?

(e) In general, let I_A and I_B be the indicators for events A and B in a probability space $(\Omega,\mathbb{P})$. What is $\mathbb{E}[I_A I_B]$, in terms of the probability of some event?

2 Marginals

(a) Can there exist three random variables X_1,X_2,X_3, each taking values in the set $\{+1,-1\}$, with the property that for every $i \neq j$, the joint distribution of X_i and X_j is given by

$$\mathbb{P}[X_i=1,X_j=-1]=\frac{1}{2} \qquad \mathbb{P}[X_i=-1,X_j=1]=\frac{1}{2} \qquad \mathbb{P}[X_i=X_j]=0? \tag{1}$$

If so, specify the joint distribution of X_1,X_2,X_3; if not, prove it.

(b) For which natural numbers $n \geq 3$ can there exist random variables $X_1, X_2, ..., X_n$, each taking values in the set $\{+1, -1\}$, with the property that for every i and j satisfying $i - j = 1 \pmod{n}$, the joint distribution of X_i and X_j is given by (1)? For any n that work, specify the joint distribution; for those that do not, prove it.

3 Testing Model Planes

Amin is testing model airplanes. He starts with n model planes which each independently have probability p of flying successfully each time they are flown, where $0 < p < 1$. Each day, he flies every single plane and keeps the ones that fly successfully (i.e. don't crash), throwing away all other models. He repeats this process for many days, where each "day" consists of Amin flying any remaining model planes and throwing away any that crash. Let X_i be the random variable representing how many model planes remain after i days. Note that $X_0 = n$. Justify your answers for each part.

(a) What is the distribution of X_1? That is, what is $\mathbb{P}[X_1 = k]$?

(b) What is the distribution of X_2? That is, what is $\mathbb{P}[X_2 = k]$? Name the distribution of X_2 and what its parameters are.

(c) Repeat the previous part for X_t for arbitrary $t \geq 1$.

(d) What is the probability that at least one model plane still remains (has not crashed yet) after t days? Do not have any summations in your answer.

(e) Considering only the first day of flights, is the event A_1 that the first and second model planes crash independent from the event B_1 that the second and third model planes crash? Recall that two events A and B are independent if $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$. Prove your answer using this definition.

(f) Considering only the first day of flights, let A_2 be the event that the first model plane crashes *and* exactly two model planes crash in total. Let B_2 be the event that the second plane crashes on the first day. What must n be equal to in terms of p such that A_2 is independent from B_2? Prove your answer using the definition of independence stated in the previous part.

(g) Are the random variables X_i and X_j, where $i < j$, independent? Recall that two random variables X and Y are independent if $\mathbb{P}[X = k_1 \cap Y = k_2] = \mathbb{P}[X = k_1]\mathbb{P}[Y = k_2]$ for all k_1 and k_2. Prove your answer using this definition.

4 Graph

Consider a random graph (undirected, no multi-edges, no self-loops) on n nodes, where each possible edge exists independently with probability p. Let X be the number of isolated nodes (nodes with degree 0).

(a) What is $\mathrm{E}(X)$? *Consider X to be the sum of the indicators X_i that vertex i is isolated. Why isn't X a binomial random variable?*

(b) (optional) What is $\mathrm{Var}(X)$?

5 Triangles in Random Graphs

Let's say we make a simple and undirected graph G on n vertices by randomly adding m edges, without replacement. In other words, we choose the first edge uniformly from all $\binom{n}{2}$ possible edges, then the second one uniformly from among the remaining $\binom{n}{2} - 1$ edges, etc. What is the expected number of triangles in G? (A triangle is a triplet of distinct vertices with all three edges present between them.)

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 12

Due: Friday 11/20 2020 at 10:00 PM
Grace period until Friday 11/20 2020 at 11:59 PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Graph

Consider a random graph (undirected, no multi-edges, no self-loops) on n nodes, where each possible edge exists independently with probability p. Let X be the number of isolated nodes (nodes with degree 0).

(a) What is $\text{Var}(X)$?

2 Whitening

Let X and Y be two random variables, with $\text{Var}(X) > 0, \text{Var}(Y) > 0$. Show that it is possible to construct $\tilde{X} = aX + bY$ and $\tilde{Y} = cX + dY$, where a, b, c, d are scalars to be chosen subject to the constraint $ad - bc \neq 0$, such that $\text{cov}(\tilde{X}, \tilde{Y}) = 0$.

You may find it unnecessary to transform Y, that is, you only need to solve for a, b to get $\text{cov}(\tilde{X}, Y) = 0$.

3 Probabilistic Bounds

A random variable X has variance $\text{Var}(X) = 9$ and expectation $\mathbb{E}[X] = 2$. Furthermore, the value of X is never greater than 10. Given this information, provide either a proof or a counterexample for the following statements.

(a) $\mathbb{E}[X^2] = 13$.

(b) $\mathbb{P}[X = 2] > 0$.

(c) $\mathbb{P}[X \geq 2] = \mathbb{P}[X \leq 2]$.

(d) $\mathbb{P}[X \leq 1] \leq 8/9$.

(e) $\mathbb{P}[X \geq 6] \leq 9/16$.

4 Subset Card Game

Jonathan and Yiming are playing a card game. Jonathan has $k > 2$ cards, and each card has a real number written on it. Jonathan tells Yiming (truthfully), that the sum of the card values is 0, and that the sum of squares of the values on the cards is 1. Specifically, if the card values are $c_1, c_2, \ldots, c_k$, then we have $\sum_{i=1}^{k} c_i = 0$ and $\sum_{i=1}^{k} c_i^2 = 1$. Jonathan and Yiming also agree on a positive target value of α.

The cards are then going to be dealt randomly in the following fashion: for each card in the deck, a fair coin is flipped. If the coin lands heads, then the card goes to Yiming, and if the coin lands tails, the card goes to Jonathan. Note that it is possible for either player to end up with no cards/all the cards.

A player wins the game if the sum of the card values in their hand is at least α, otherwise it is a tie.

(a) Prove that the probability that Yiming wins is at most $\frac{1}{8\alpha^2}$.

(b) Find a deck of k cards and target value α where the probability that Yiming wins is exactly $\frac{1}{8\alpha^2}$.

5 Just One Tail, Please

Let X be some random variable with finite mean and variance which is not necessarily non-negative. The *extended* version of Markov's Inequality states that for a non-negative function $\phi(x)$ which is monotonically increasing for $x > 0$ and some constant $\alpha > 0$,

$$\mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[\phi(X)]}{\phi(\alpha)}$$

Suppose $\mathbb{E}[X] = 0$, $\text{Var}(X) = \sigma^2 < \infty$, and $\alpha > 0$.

(a) Use the extended version of Markov's Inequality stated above with $\phi(x) = (x+c)^2$, where c is some positive constant, to show that:

$$\mathbb{P}(X \geq \alpha) \leq \frac{\sigma^2 + c^2}{(\alpha + c)^2}$$

(b) Note that the above bound applies for all positive c, so we can choose a value of c to minimize the expression, yielding the best possible bound. Find the value for c which will minimize the RHS expression (you may assume that the expression has a unique minimum). Plug in the minimizing value of c to prove the following bound:

$$\mathbb{P}(X \geq \alpha) \leq \frac{\sigma^2}{\alpha^2 + \sigma^2}.$$

(c) Recall that Chebyshev's inequality provides a two-sided bound. That is, it provides a bound on $\mathbb{P}(|X - \mathbb{E}[X]| \geq \alpha) = \mathbb{P}(X \geq \mathbb{E}[X] + \alpha) + \mathbb{P}(X \leq \mathbb{E}[X] - \alpha)$. If we only wanted to bound the probability of one of the tails, e.g. if we wanted to bound $\mathbb{P}(X \geq \mathbb{E}[X] + \alpha)$, it is tempting to just divide the bound we get from Chebyshev's by two. Why is this not always correct in general? Provide an example of a random variable X (does not have to be zero-mean) and a constant α such that using this method (dividing by two to bound one tail) is not correct, that is, $\mathbb{P}(X \geq \mathbb{E}[X] + \alpha) > \frac{\text{Var}(X)}{2\alpha^2}$ or $\mathbb{P}(X \leq \mathbb{E}[X] - \alpha) > \frac{\text{Var}(X)}{2\alpha^2}$.

Now we see the use of the bound proven in part (b) - it allows us to bound just one tail while still taking variance into account, and does not require us to assume any property of the random variable. Note that the bound is also always guaranteed to be less than 1 (and therefore at least somewhat useful), unlike Markov's and Chebyshev's inequality!

(d) Let's try out our new bound on a simple example. Suppose X is a positively-valued random variable with $\mathbb{E}[X] = 3$ and $\text{Var}(X) = 2$. What bound would Markov's inequality give for $\mathbb{P}[X \geq 5]$? What bound would Chebyshev's inequality give for $\mathbb{P}[X \geq 5]$? What about for the bound we proved in part (b)? (*Note*: Recall that the bound from part (b) only applies for zero-mean random variables.)

6 Sum of Poisson Variables

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.

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

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 13

Due: Monday 11/30 2020 at 10:00PM
Grace period until Monday 11/30 2020 at 11:59PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Random Cuckoo Hashing

Cuckoo birds are parasitic beasts. They are known for hijacking the nests of other bird species and evicting the eggs already inside. Cuckoo hashing is inspired by this behavior. In cuckoo hashing, when we get a collision, the element that was already there gets evicted and rehashed.

We study a simple (but ineffective, as we'll see) version of cuckoo hashing, where all hashes are random. Let's say we want to hash n pieces of data $d_1, d_2, \ldots, d_n$ into n possible hash buckets labeled $1, \ldots, n$. We hash the $d_1, \ldots, d_n$ in that order. When hashing d_i, we assign it a random bucket chosen uniformly from $1, \ldots, n$. If there is no collision, then we place d_i into that bucket. If there is a collision with some other d_j, we evict d_j and assign it another random bucket uniformly from $1, \ldots, n$. (It is possible that d_j gets assigned back to the bucket it was just evicted from!) We again perform the eviction step if we get another collision. We keep doing this until there is no more collision, and we then introduce the next piece of data, d_{i+1} to the hash table.

(a) What is the probability that there are no collisions over the entire process of hashing $d_1, \ldots, d_n$ to buckets $1, \ldots, n$? What value does the probability tend towards as n grows very large?

(b) Assume we have already hashed $d_1, \ldots, d_{n-1}$, and they each occupy their own bucket. We now introduce d_n into our hash table. What is the expected number of collisions that we'll see while hashing d_n? (*Hint*: What happens when we hash d_n and get a collision, so we evict some other d_i and have to hash d_i? Are we at a situation that we've seen before?)

(c) Generalize the previous part: Assume we have already hashed $d_1, \ldots, d_{k-1}$ successully, where $1 \leq k \leq n$. Let C_k be the number of collisions that we'll see while hashing d_k. What is $\mathbb{E}[C_k]$?

(d) Let C be the total number of collisions over the entire process of hashing $d_1, \ldots, d_n$. What is $\mathbb{E}[C]$? You may leave your answer as a summation.

2 Geometric and Poisson

Let $X \sim \text{Geo}(p)$ and $Y \sim \text{Poisson}(\lambda)$ be independent. random variables. Compute $\mathbb{P}(X > Y)$. Your final answer should not have summations.

3 Exploring the Geometric Distribution

Suppose $X \sim \text{Geometric}(p)$ and $Y \sim \text{Geometric}(q)$ are independent. Find the distribution of $\min\{X, Y\}$ and justify your answer.

4 Lunch Meeting

Alice and Bob agree to try to meet for lunch between 12 PM and 1 PM at their favorite sushi restaurant. Being extremely busy, they are unable to specify their arrival times exactly, and can say only that each of them will arrive (independently) at a time that is uniformly distributed within the hour. In order to avoid wasting precious time, if the other person is not there when they arrive they agree to wait exactly fifteen minutes before leaving. What is the probability that they will actually meet for lunch? (hint: Sketch the joint distribution of the arrival times of Alice and Bob. What parts of the distribution corresponds to them meeting for lunch?)

CS 70 Discrete Mathematics and Probability Theory
Fall 2020

HW 14

Due: Sunday 12/06 2020 at 10:00 PM
Grace period until Sunday 12/06 2020 at 11:59PM

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Short Answer

(a) Let X be uniform on the interval $[0,2]$, and define $Y = 2X+1$. Find the PDF, CDF, expectation, and variance of Y.

(b) Let X and Y have joint distribution

$$f(x,y) = \begin{cases} cxy + 1/4 & x \in [1,2] \text{ and } y \in [0,2] \\ 0 & \text{else} \end{cases}$$

Find the constant c. Are X and Y independent?

2 Continuous Probability Continued

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

(a) Assume $\text{Bob}_1, \text{Bob}_2, \ldots, \text{Bob}_k$ 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)?

3 Exponential Distributions: Lightbulbs

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?

4 Useful Uniforms

Let X be a continuous random variable whose image is all of $\mathbb{R}$; that is, $\mathbb{P}[X \in (a,b)] > 0$ for all $a, b \in \mathbb{R}$ and $a \neq b$.

(a) Give an example of a distribution that X could have, and one that it could not.

(b) Show that the CDF F of X is strictly increasing. That is, $F(x+\varepsilon) > F(x)$ for any $\varepsilon > 0$. Argue why this implies that $F : \mathbb{R} \to (0,1)$ must be invertible.

(c) Let U be a uniform random variable on $(0,1)$. What is the distribution of $F^{-1}(U)$?

(d) Your work in part (c) shows that in order to sample X, it is enough to be able to sample U. If X was a discrete random variable instead, taking finitely many values, can we still use U to sample X?

5 Uniform Means

To keep the doctor away, Bob goes to the supermarket to buy an apple. Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed uniform random variables on the interval $[0,1]$ (where n is a positive integer), where X_i is the quality of the ith apple Bob sees.

(a) Let $Y = \min\{X_1, X_2, \ldots, X_n\}$ be the quality of the worst apple Bob will see. Find $\mathbb{E}(Y)$. [*Hint*: Use the tail sum formula, which says the expected value of a nonnegative random variable is $\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > x)\,\mathrm{d}x$. Note that we can use the tail sum formula since $Y \geq 0$.]

(b) Let $Z = \max\{X_1, X_2, \ldots, X_n\}$ be the quality of the best apple Bob will see. Find $\mathbb{E}(Z)$. [*Hint*: Find the CDF.]

6 Darts but with ML

Suppose Alice and Bob are playing darts on a circular board with radius 1. When Alice throws a dart, the distance of the dart from the center is uniform $[0,1]$. When Bob throws the dart, the location of the dart is uniform over the whole board. Let X be the random variable corresponding to the distance of the player's dart from the center of the board.

(a) What is the pdf of X if Alice throws

(b) What is the pdf of X if Bob throws

(c) Suppose we let Alice throw the dart with probability p, and let Bob throw otherwise. What is the pdf of X (your answer should be in terms of p)?

(d) Using the same premise as in part c, suppose you observe a dart on the board but don't know who threw it. Let x be the dart's distance from the center. We would like to come up with a decision rule to determine whether Alice or Bob is more likely to have thrown the dart given your observation, x. Specifically, if we let A be the event that Alice threw the dart and B be the event that Bob threw, we want to guess A if $\mathbb{P}[A|X \in [x,x+dx]] > \mathbb{P}[B|X \in [x,x+dx]]$ (what do these two probabilities have to sum up to?). For what values of x would we guess A? (your answer should be in terms of p)

7 Sampling a Gaussian With Uniform

In this question, we will see one way to generate a normal random variable if we have access to a random number generator that outputs numbers between 0 and 1 uniformly at random.

As a general comment, remember that showing two random variables have the same CDF or PDF is sufficient for showing that they have the same distribution.

(a) First, let us see how to generate an exponential random variable with a uniform random variable. Let $U_1 \sim Uniform(0,1)$. Prove that $-\ln U_1 \sim Expo(1)$.

(b) Let $N_1, N_2 \sim \mathcal{N}(0,1)$, where N_1 and N_2 are independent. Prove that $N_1^2 + N_2^2 \sim Expo(1/2)$.

Hint: You may use the fact that over a region R, if we convert to polar coordinates $(x,y) \to (r,\theta)$, then the double integral over the region R will be

$$\iint_R f(x,y)\,dx\,dy = \iint_R f(r\cos\theta, r\sin\theta)\cdot r\,dr\,d\theta.$$

(c) Suppose we have two uniform random variables, U_1 and U_2. How would you transform these two random variables into a normal random variable with mean 0 and variance 1?

Hint: What part (b) tells us is that the point (N_1,N_2) will have a distance from the origin that is distributed as the square root of an exponential distribution. Try to use U_1 to sample the radius, and then use U_2 to sample the angle.

CS 70 Discrete Mathematics and Probability Theory
Fall 2020 HW 15

Due: n/a
Grace period until n/a

Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.)

1 Playing Blackjack

You are playing a game of Blackjack where you start with \$100. You are a particularly risk-loving player who does not believe in leaving the table until you either make \$400, or lose all your money. At each turn you either win \$100 with probability p, or you lose \$100 with probability $1-p$.

(a) Formulate this problem as a Markov chain i.e. define your state space, transition probabilities, and determine your starting state.

(b) Find the probability that you end the game with \$400.

2 Markov's Coupon Collecting

Courtney is home for Thanksgiving and needs to make some trips to the Traitor Goes grocery store to prepare for the big turkey feast. Each time she goes to the store before the holiday, she receives one of the n different coupons that are being given away. You may recall that we studied how to calculate the expected number of trips to the store needed to collect at least one of each coupon. Using geometric distributions and indicator variables, we determined that expected number of trips to be $n(\ln n + \gamma)$.

Let's re-derive that, this time with a Markov chain model and first-step equations.

(a) Define the states and transition probabilities for each state (explain what states can be transitioned to, and what probabilities those transitions occur with).

(b) Now set up first-step equations and solve for the expected number of grocery store trips Courtney needs to make before Thanksgiving so that she can have at least one of each of the n distinct coupons.

3 Reflecting Random Walk

Alice starts at vertex 0 and wishes to get to vertex n. When she is at vertex 0 she has a probability of 1 of transitioning to vertex 1. For any other vertex i, there is a probability of $1/2$ of transitioning to $i+1$ and a probability of $1/2$ of transitioning to $i-1$.

(a) What is the expected number of steps Alice takes to reach vertex n? Write down the hitting-time equations, but do not solve them yet.

(b) Solve the hitting-time equations. [*Hint*: Let R_i denote the expected number of steps to reach vertex n starting from vertex i. As a suggestion, try writing R_0 in terms of R_1; then, use this to express R_1 in terms of R_2; and then use this to express R_2 in terms of R_3, and so on. See if you can notice a pattern.]

4 Boba in a Straw

Imagine that Jonathan is drinking milk tea and he has a very short straw: it has enough room to fit two boba (see figure).

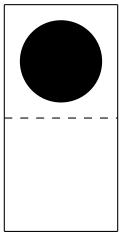


Figure 1: A straw with one boba currently inside. The straw only has enough room to fit two boba.

Here is a formal description of the drinking process: We model the straw as having two "components" (the top component and the bottom component). At any given time, a component can contain nothing, or one boba. As Jonathan drinks from the straw, the following happens every second:

1. The contents of the top component enter Jonathan's mouth.
2. The contents of the bottom component move to the top component.
3. With probability p, a new boba enters the bottom component; otherwise the bottom component is now empty.

Help Jonathan evaluate the consequences of his incessant drinking!

(a) At the very start, the straw starts off completely empty. What is the expected number of seconds that elapse before the straw is completely filled with boba for the first time? [Write down the equations; you do not have to solve them.]

(b) Consider a slight variant of the previous part: now the straw is narrower at the bottom than at the top. This affects the drinking speed: if either (i) a new boba is about to enter the bottom component or (ii) a boba from the bottom component is about to move to the top component, then the action takes two seconds. If both (i) and (ii) are about to happen, then the action takes three seconds. Otherwise, the action takes one second. Under these conditions, answer the previous part again. [Write down the equations; you do not have to solve them.]

(c) Jonathan was annoyed by the straw so he bought a fresh new straw (the straw is no longer narrow at the bottom). What is the long-run average rate of Jonathan's calorie consumption? (Each boba is roughly 10 calories.)

(d) What is the long-run average number of boba which can be found inside the straw? [Maybe you should first think about the long-run distribution of the number of boba.]