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

DIS 0A

1 Propositional Practice

Note 1

Convert the following English sentences into propositional logic and the following propositions into English. State whether or not each statement is true with brief justification.

(a) There is a real number which is not rational.

(b) All integers are natural numbers or are negative, but not both.

(c) If a natural number is divisible by 6, it is divisible by 2 or it is divisible by 3.

(d) $(\forall x \in \mathbb{Z})\ (x \in \mathbb{Q})$

(e) $(\forall x \in \mathbb{Z})\ (((2 \mid x) \lor (3 \mid x)) \implies (6 \mid x))$

(f) $(\forall x \in \mathbb{N})\ ((x > 7) \implies ((\exists a, b \in \mathbb{N})\ (a + b = x)))$

2 Truth Tables

Note 1

Determine whether the following equivalences hold, by writing out truth tables. Clearly state whether or not each pair is equivalent.

(a) $P \land (Q \lor P) \equiv P \land Q$

(b) $(P \lor Q) \land R \equiv (P \land R) \lor (Q \land R)$

(c) $(P \land Q) \lor R \equiv (P \lor R) \land (Q \lor R)$

3 Implication

Note 0
Note 1

Which of the following implications are always true, regardless of P? Give a counterexample for each false assertion (i.e. come up with a statement $P(x,y)$ that would make the implication false).

(a) $\forall x \forall y P(x,y) \implies \forall y \forall x P(x,y)$.

(b) $\forall x \exists y P(x,y) \implies \exists y \forall x P(x,y)$.

(c) $\exists x \forall y P(x,y) \implies \forall y \exists x P(x,y)$.

CS 70 Spring 2024

Discrete Mathematics and Probability Theory
Seshia, Sinclair

DIS 0B

1 Perfect Square

Note 2

(a) Prove that if n^2 is odd, then n must also be odd.

(b) Prove that if n^2 is odd, then n^2 can be written in the form $8k+1$ for some integer k.

2 Numbers of Friends

Note 2

Prove that if there are $n \geq 2$ people at a party, then at least 2 of them have the same number of friends at the party. Assume that friendships are always reciprocated: that is, if Alice is friends with Bob, then Bob is also friends with Alice.

(Hint: The Pigeonhole Principle states that if n items are placed in m containers, where $n > m$, at least one container must contain more than one item. You may use this without proof.)

3 Pebbles

Note 2

Suppose you have a rectangular array of pebbles, where each pebble is either red or blue. Suppose that for every way of choosing one pebble from each column, there exists a red pebble among the chosen ones.

Prove that there must exist an all-red column.

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair
DIS 1A

1 Natural Induction on Inequality

Note 3

Prove that if $n \in \mathbb{N}$ and $x > 0$, then $(1+x)^n \geq 1+nx$.

2 Make It Stronger

Note 3

Suppose that the sequence $a_1, a_2, \ldots$ is defined by $a_1 = 1$ and $a_{n+1} = 3a_n^2$ for $n \geq 1$. We want to prove that

$$a_n \leq 3^{(2^n)}$$

for every positive integer n.

(a) Suppose that we want to prove this statement using induction. Can we let our inductive hypothesis be simply $a_n \leq 3^{(2^n)}$? Attempt an induction proof with this hypothesis to show why this does not work.

(b) Try to instead prove the statement $a_n \leq 3^{(2^n-1)}$ using induction.

(c) Why does the hypothesis in part (b) imply the overall claim?

3 Binary Numbers

Note 3

Prove that every positive integer n can be written in binary. In other words, prove that for any positive integer n, we can write

$$n = c_k \cdot 2^k + c_{k-1} \cdot 2^{k-1} + \cdots + c_1 \cdot 2^1 + c_0 \cdot 2^0,$$

for some $k \in \mathbb{N}$ and $c_i \in \{0,1\}$ for all $i \leq k$.

4 Fibonacci for Home

Note 3

Recall, the Fibonacci numbers, defined recursively as

$F_1 = 1$, $F_2 = 1$, and $F_n = F_{n-2} + F_{n-1}$.

Prove that every third Fibonacci number is even. For example, $F_3 = 2$ is even and $F_6 = 8$ is even.

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair DIS 1B

1 Stable Matching

Note 4

Consider the set of jobs $J = \{1, 2, 3\}$ and the set of candidates $C = \{A, B, C\}$ with the following preferences.

Jobs	Candidates
1	A > B > C
2	B > A > C
3	A > B > C

Candidates	Jobs
A	2 > 1 > 3
B	1 > 3 > 2
C	1 > 2 > 3

Run the traditional propose-and-reject algorithm on this example. How many days does it take and what is the resulting pairing? (Show your work.)

2 Propose-and-Reject Proofs

Note 4

Prove the following statements about the traditional propose-and-reject algorithm.

(a) In any execution of the algorithm, if a candidate receives a proposal on day i, then she receives some proposal on every day thereafter until termination.

(b) In any execution of the algorithm, if a candidate receives no proposal on day i, then she receives no proposal on any previous day j, $1 \leq j < i$.

(c) In any execution of the algorithm, there is at least one candidate who only receives a single proposal. (Hint: use the parts above!)

3 Be a Judge

Note 4

By stable matching instance, we mean a set of jobs and candidates and their preference lists. For each of the following statements, indicate whether the statement is True or False and justify your answer with a short 2-3 line explanation:

(a) There is a stable matching instance for n jobs and n candidates for $n > 1$, such that in a stable matching algorithm with jobs proposing, every job ends up with its least preferred candidate.

(b) In a stable matching instance, if job J and candidate C each put each other at the top of their respective preference lists, then J must be paired with C in every stable pairing.

(c) In a stable matching instance with at least two jobs and two candidates, if job J and candidate C each put each other at the bottom of their respective preference lists, then J cannot be paired with C in any stable pairing.

(d) For every $n > 1$, there is a stable matching instance for n jobs and n candidates which has an **unstable** pairing where **every** unmatched job-candidate pair is a rogue couple or pairing.

4 Stable Matching III

Note 4

(a) True or False?

(i) If a candidate accidentally rejects a job she prefers on a given day, then the algorithm still always ends with a matching.

(ii) The Propose-and-Reject Algorithm never produces a candidate-optimal matching.

(iii) If the same job is last on the preference list of every candidate, the job must end up with its least preferred candidate.

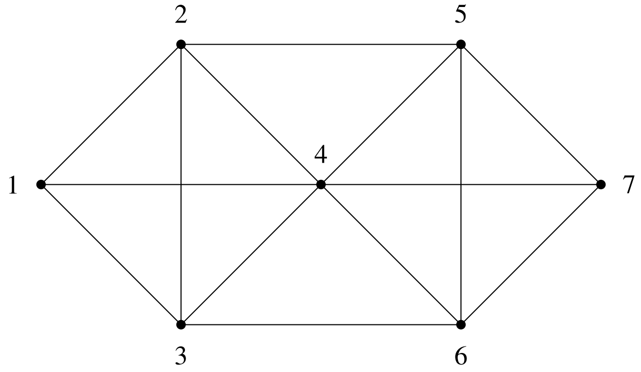
(b) As you've seen from lecture, the jobs-proposing Propose-and-Reject Algorithm produces an employer-optimal stable matching. Let's see if the candidate have any way of improving their standing. Suppose exactly one of the candidates has the power to arbitrarily reject one proposal, regardless of which job she has on her string (if any). Construct an example that illustrates the following: for any $n \geq 2$, there exists a stable matching instance for which using this power helps **every** candidate, i.e. every candidate gets a better job than she would have gotten under the jobs-proposing Propose-and-Reject Algorithm.

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

DIS 2A

1 Eulerian Tour and Eulerian Walk

Note 5



(a) Is there an Eulerian tour in the graph above? If no, give justification. If yes, provide an example.

(b) Is there an Eulerian walk in the graph above? An Eulerian walk is a walk that uses each edge exactly once. If no, give justification. If yes, provide an example.

(c) What is the condition that there is an Eulerian walk in an undirected graph? Briefly justify your answer.

2 Coloring Trees

Note 5

(a) Prove that all trees with at least 2 vertices have at least two leaves. Recall that a leaf is defined as a node in a tree with degree exactly 1.

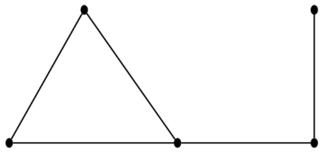
(b) Prove that all trees with at least 2 vertices are *bipartite*: the vertices can be partitioned into two groups so that every edge goes between the two groups.

[*Hint:* Use induction on the number of vertices.]

3 Degree Sequences

Note 5

The *degree sequence* of a graph is the sequence of the degrees of the vertices, arranged in descending order, with repetitions as needed. For example, the degree sequence of the following graph is $(3,2,2,2,1)$.



For each of the parts below, determine if there exists a simple undirected graph G (i.e. a graph without self-loops and multiple-edges) having the given degree sequence. Justify your claim.

(a) $(3,3,2,2)$

(b) $(3,2,2,2,2,1,1)$

(c) $(6,2,2,2)$

(d) $(4,4,3,2,1)$

CS 70 Spring 2024 — Discrete Mathematics and Probability Theory — Seshia, Sinclair — DIS 2B

1 Short Answers

Note 5

In each part below, provide the number/equation and brief justification.

(a) A connected planar simple graph has 5 more edges than it has vertices. How many faces does it have?

(b) How many edges need to be removed from a 3-dimensional hypercube to get a tree?

(c) The Euler's formula $v - e + f = 2$ requires the planar graph to be connected. What is the analogous formula for planar graphs wth k connected components?

2 Always, Sometimes, or Never

Note 5

In each part below, you are given some information about a graph G. Using only the information in the current part, say whether G 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) G can be vertex-colored with 4 colors.

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

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

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

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

3 Graph Coloring

Note 5

Prove that a graph with maximum degree at most k is $(k+1)$-colorable.

4 Hypercubes

Note 5

The vertex set of the n-dimensional hypercube $G = (V,E)$ is given by $V = \{0,1\}^n$ (recall that $\{0,1\}^n$ denotes the set of all n-bit strings). There is an edge between two vertices x and y if and only if x and y differ in exactly one bit position.

(a) Draw 1-, 2-, and 3-dimensional hypercubes and label the vertices using the corresponding bit strings.

(b) Show that the edges of an n-dimensional hypercube can be colored using n colors so that no pair of edges sharing a common vertex have the same color.

(c) Show that for any $n \geq 1$, the n-dimensional hypercube is bipartite.

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 3A**

1 Party Tricks

Note 6

You are at a party celebrating your completion of the CS 70 midterm. Show off your modular arithmetic skills and impress your friends by quickly figuring out the last digit(s) of each of the following numbers:

(a) Find the last digit of 11^{3142}.

(b) Find the last digit of 9^{9999}.

(c) Find the last digit of 3^{641}.

2 Modular Potpourri

Note 6

Prove or disprove the following statements:

(a) There exists some $x \in \mathbb{Z}$ such that $x \equiv 3 \pmod{16}$ and $x \equiv 4 \pmod{6}$.

(b) $2x \equiv 4 \pmod{12} \iff x \equiv 2 \pmod{12}$.

(c) $2x \equiv 4 \pmod{12} \iff x \equiv 2 \pmod{6}$.

3 Modular Inverses

Note 6

Recall the definition of inverses from lecture: let $a, m \in \mathbb{Z}$ and $m > 0$; if $x \in \mathbb{Z}$ satisfies $ax \equiv 1 \pmod{m}$, then we say x is an **inverse of** a **modulo** m.

Now, we will investigate the existence and uniqueness of inverses.

(a) Is 3 an inverse of 5 modulo 10?

(b) Is 3 an inverse of 5 modulo 14?

(c) For all $n \in \mathbb{N}$, is $3 + 14n$ an inverse of 5 modulo 14?

(d) Does 4 have an inverse modulo 8?

(e) Suppose $x, x' \in \mathbb{Z}$ are both inverses of a modulo m. Is it possible that $x \not\equiv x' \pmod{m}$?

4 Fibonacci GCD

Note 6

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 $n \geq 1$, $\gcd(F_n, F_{n-1}) = 1$.

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 3B**

1 Extended Euclid: Two Ways

Note 6

In this problem, we will explore the Extended Euclid's Algorithm: first, the traditional implementation, and second, a faster, iterative version. Both ways yield the same result.

Parts (b) and (c) explore the traditional Extended Euclid's Algorithm. The bolded numbers below keep track of which numbers appeared as inputs to the gcd call. Remember that we are interested in writing the GCD as a linear combination of the original inputs, so we don't want to accidentally simplify the expressions and eliminate the inputs.

(a) As motivation, suppose we've found values of a and b such that $54a + 17b = 1$. With this knowledge, what is $17^{-1} \pmod{54}$?

(b) Note that x mod y, by definition, is always x minus a multiple of y. So, in the execution of Euclid's algorithm, each newly introduced value can always be expressed as a "combination" of the previous two, like so:

$$\begin{aligned}
\gcd(54,17) &= \gcd(17,3) & \mathbf{3} &= 1 \times \mathbf{54} - 3 \times \mathbf{17} \\
&= \gcd(3,2) & \mathbf{2} &= 1 \times \mathbf{17} - \_\_\_\_ \times \mathbf{3} \\
&= \gcd(2,1) & \mathbf{1} &= 1 \times \mathbf{3} - \_\_\_\_ \times \mathbf{2} \\
&= \gcd(1,0) & [\mathbf{0} &= 1 \times \mathbf{2} - \_\_\_\_ \times \mathbf{1}] \\
&= 1. & &
\end{aligned}$$

(Fill in the blanks)

(c) Recall that our goal is to fill out the blanks in

$$1 = \_\_\_\_ \times \mathbf{54} + \_\_\_\_ \times \mathbf{17}.$$

To do so, we work back up from the bottom, and express the gcd above as a combination of the two arguments on each of the previous lines:

$$\begin{aligned}
1 &= \_\_\_\_ \times \mathbf{3} + \_\_\_\_ \times \mathbf{2} \\
&= \\
&= \_\_\_\_ \times \mathbf{17} + \_\_\_\_ \times \mathbf{3} \\
&= \\
&= \_\_\_\_ \times \mathbf{54} + \_\_\_\_ \times \mathbf{17}
\end{aligned}$$

What does this imply, in this case, about the multiplicative inverse of 17, in arithmetic mod 54?

(d) In the previous parts, we used a recursive method to write $\gcd(54,17)$ as a linear combination of 54 and 17. We can also compute the same result iteratively—this is an alternative to the above method that is oftentimes faster. We begin by writing equations for our initial arguments, 54 and 17, as a linear combination of themselves:

$$
\begin{aligned}
54 &= 1 \times \mathbf{54} + 0 \times \mathbf{17} && (E_1) \\
17 &= 0 \times \mathbf{54} + 1 \times \mathbf{17} && (E_2)
\end{aligned}
$$

We can then use these initial equations (labeled E_1 and E_2 for ease of reference) to iteratively write reduced values as linear combinations of 54 and 17, until we are able to write an equation for $\gcd(54,17)$, as desired.

In particular, we want to subtract as many multiples of the second equation as possible from the first to create a new equation with a lower LHS value. We can keep iterating until the LHS becomes $\gcd(54,17) = 1$.

$$
\begin{aligned}
\_\_\_\_ &= \_\_\_\_ \times \mathbf{54} + \_\_\_\_ \times \mathbf{17} && (E_3 = E_1 - \_\_\_\_ \times E_2) \\
\_\_\_\_ &= \_\_\_\_ \times \mathbf{54} + \_\_\_\_ \times \mathbf{17} && (E_4 = E_2 - \_\_\_\_ \times E_3) \\
1 &= \_\_\_\_ \times \mathbf{54} + \_\_\_\_ \times \mathbf{17} && (E_5 = E_3 - \_\_\_\_ \times E_4)
\end{aligned}
$$

What does this imply, in this case, about the multiplicative inverse of 17, in arithmetic mod 54? Verify that your answer is equivalent to the previous part.

(e) Calculate the gcd of 17 and 39, and determine how to express this as a “combination” of 17 and 39. What does this imply, in this case, about the multiplicative inverse of 17, in arithmetic mod 39?

2 Chinese Remainder Theorem Practice

Note 6

In this question, you will solve for a natural number x such that,

$$\begin{aligned} x &\equiv 1 \pmod{3} \\ x &\equiv 3 \pmod{7} \\ x &\equiv 4 \pmod{11} \end{aligned} \tag{1}$$

(a) Suppose you find 3 natural numbers a, b, c that satisfy the following properties:

$$a \equiv 1 \pmod{3}\ ;\ a \equiv 0 \pmod{7}\ ;\ a \equiv 0 \pmod{11}, \tag{2}$$

$$b \equiv 0 \pmod{3}\ ;\ b \equiv 1 \pmod{7}\ ;\ b \equiv 0 \pmod{11}, \tag{3}$$

$$c \equiv 0 \pmod{3}\ ;\ c \equiv 0 \pmod{7}\ ;\ c \equiv 1 \pmod{11}. \tag{4}$$

Show how you can use the knowledge of a, b and c to compute an x that satisfies (1).

In the following parts, you will compute natural numbers a, b and c that satisfy the above 3 conditions and use them to find an x that satisfies (1).

(b) Find a natural number a that satisfies (2). That is, $a \equiv 1 \pmod{3}$ and is a multiple of 7 and 11.

It may help to start with a number that is a multiple of both 7 and 11; what number should we multiply this by in order to make it equivalent to $1 \pmod{3}$?

(c) Find a natural number b that satisfies (3). That is, $b \equiv 1 \pmod{7}$ and is a multiple of 3 and 11.

(d) Find a natural number c that satisfies (4). That is, $c \equiv 1 \pmod{11}$ and is a multiple of 3 and 7.

(e) Putting together your answers for parts (a), (b), (c) and (d), report an x that satisfies (1).

3 Baby Fermat

Note 6

Assume that a does have a multiplicative inverse mod m. Let us prove that its multiplicative inverse can be written as $a^k \pmod{m}$ for some $k \geq 0$.

(a) Consider the infinite sequence $a, a^2, a^3, \ldots \pmod{m}$. Prove that this sequence has repetitions.

(**Hint:** Consider the Pigeonhole Principle.)

(b) Assuming that $a^i \equiv a^j \pmod{m}$, where $i > j$, what is the value of $a^{i-j} \pmod{m}$?

(c) Prove that the multiplicative inverse can be written as $a^k \pmod{m}$. What is k in terms of i and j?

1 RSA Warm-Up

Note 7

Consider an RSA scheme with modulus $N = pq$, where p and q are distinct prime numbers larger than 3.

(a) What is wrong with using the exponent $e = 2$ in an RSA public key?

(b) Recall that e must be relatively prime to $p-1$ and $q-1$. Find a condition on p and q such that $e = 3$ is a valid exponent.

(c) Now suppose that $p = 5$, $q = 17$, and $e = 3$. What is the public key?

(d) What is the private key?

(e) Alice wants to send a message $x = 10$ to Bob. What is the encrypted message $E(x)$ she sends using the public key?

(f) Suppose Bob receives the message $y = 19$ from Alice. What equation would he use to decrypt the message? What is the decrypted message?

2 RSA with Multiple Keys

Note 7

Members of a secret society know a secret word. They transmit this secret word x between each other many times, each time encrypting it with the RSA method. Eve, who is listening to all of their communications, notices that in all of the public keys they use, the exponent e is the same. Therefore the public keys used look like $(N_1, e), \ldots, (N_k, e)$ where no two N_i's are the same. Assume that the message is x such that $0 \le x < N_i$ for every i.

Further, in all of the subparts, you may assume that Eve knows the details of the modified RSA schemes (i.e. Eve knows the format of the N_i's, but not the specific values used to compute the N_i's).

(a) Suppose Eve sees the public keys $(p_1 q_1, 7)$ and $(p_1 q_2, 7)$ as well as the corresponding transmissions. Can Eve use this knowledge to break the encryption? If so, how? Assume that Eve cannot compute prime factors efficiently. Think of p_1, q_1, q_2 as massive 1024-bit numbers. Assume p_1, q_1, q_2 are all distinct and are valid primes for RSA to be carried out.

(b) The secret society has wised up to Eve and changed their choices of N, in addition to changing their word x. Now, Eve sees keys $(p_1 q_1, 3)$, $(p_2 q_2, 3)$, and $(p_3 q_3, 3)$ along with their transmissions. Argue why Eve cannot break the encryption in the same way as above. Assume $p_1, p_2, p_3, q_1, q_2, q_3$ are all distinct and are valid primes for RSA to be carried out.

(c) Let's say the secret x was not changed ($e = 3$), so they used the same public keys as before, but did not transmit different messages. How can Eve figure out x?

3 RSA for Concert Tickets

Note 7

Alice wants to tell Bob her concert ticket number, m, which is an integer between 0 and 100 inclusive. She wants to tell Bob over an insecure channel that Eve can listen in on, but Alice does not want Eve to know her ticket number.

(a) Bob announces his public key $(N = pq, e)$, where N is large (512 bits). Alice encrypts her message using RSA. Eve sees the encrypted message, and figures out what Alice's ticket number is. How did she do it?

(b) Alice decides to be a bit more elaborate. She picks a random number r that is 256 bits long, so that it is too hard to guess. She encrypts that and sends it to Bob, and also computes rm, encrypts that, and sends it to Bob. Eve is aware of what Alice did, but does not know the value of r. How can she figure out m? (You may assume that r is coprime to N.)

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

DIS 4B

1 Polynomial Practice

Note 8

(a) If f and g are non-zero real polynomials, how many real roots do the following polynomials have at least? How many can they have at most? (Your answer may depend on the degrees of f and g.)

(i) $f+g$

(ii) $f \cdot g$

(iii) f/g, assuming that f/g is a polynomial

(b) Now let f and g be polynomials over $\mathrm{GF}(p)$.

(i) We say a polynomial $f = 0$ if $\forall x, f(x) = 0$. Show that if $f \cdot g = 0$, it is not always true that either $f = 0$ or $g = 0$.

(ii) How many f of degree *exactly* $d < p$ are there such that $f(0) = a$ for some fixed $a \in \{0, 1, \dots, p-1\}$?

(c) Find a polynomial f over $\mathrm{GF}(5)$ that satisfies $f(0) = 1, f(2) = 2, f(4) = 0$. How many such polynomials of degree at most 4 are there?

2 Lagrange Interpolation in Finite Fields

Note 8

Find a unique polynomial $p(x)$ of degree at most 2 that passes through points $(-1,3)$, $(0,1)$, and $(1,2)$ in modulo 5 arithmetic using the Lagrange interpolation.

(a) Find $p_{-1}(x)$ where $p_{-1}(0) \equiv p_{-1}(1) \equiv 0 \pmod{5}$ and $p_{-1}(-1) \equiv 1 \pmod{5}$.

(b) Find $p_0(x)$ where $p_0(-1) \equiv p_0(1) \equiv 0 \pmod{5}$ and $p_0(0) \equiv 1 \pmod{5}$.

(c) Find $p_1(x)$ where $p_1(-1) \equiv p_1(0) \equiv 0 \pmod{5}$ and $p_1(1) \equiv 1 \pmod{5}$.

(d) Construct $p(x)$ using a linear combination of $p_{-1}(x)$, $p_0(x)$, and $p_1(x)$.

3 Secrets in the United Nations

Note 8

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.

4 To The Moon!

Note 8

A secret number s is required to launch a rocket, and Alice distributed the values $(1,p(1)),(2,p(2)),\ldots,(n+1,p(n+1))$ of a degree n polynomial p to a group of \$GME holders $\text{Bob}_1,\ldots,\text{Bob}_{n+1}$. As usual, she chose p such that $p(0)=s$. Bob_1 through Bob_{n+1} now gather to jointly discover the secret. However, Bob_1 is secretly a partner at Melvin Capital and already knows s, and wants to sabotage $\text{Bob}_2,\ldots,\text{Bob}_{n+1}$, making them believe that the secret is in fact some fixed $s' \neq s$. How could he achieve this? In other words, what value should he report (in terms variables known in the problem, such as s',s or y_1) in order to make the others believe that the secret is s'?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair DIS 5A

1 Berlekamp-Welch Warm Up

Note 8
Note 9

Let $P(i)$, a polynomial applied to the input i, be the original encoded polynomial before sent, and let r_i be the received info for the input i which may or may not be corrupted.

(a) If you want to send a length-n message, what should the degree of $P(x)$ be? Why?

(b) When does $r_i = P(i)$? When does r_i not equal $P(i)$?

(c) If there are at most k erasure errors, how many packets should you send? If there are at most k general errors, how many packets should you send? (We will see the reason for this later.) Now we will only consider general errors.

(d) What do the roots of the error polynomial $E(x)$ represent? Does the receiver know the roots of $E(x)$? If there are at most k errors, what is the maximum degree of $E(x)$? Using the information about the degree of $P(x)$ and $E(x)$, what is the degree of $Q(x) = P(x)E(x)$?

(e) Why is the equation $Q(i) = P(i)E(i) = r_iE(i)$ always true? (Consider what happens when $P(i) = r_i$, and what happens when $P(i)$ does not equal r_i.)

(f) In the polynomials $Q(x)$ and $E(x)$, how many total unknown coefficients are there? (These are the variables you must solve for. Think about the degree of the polynomials.) When you receive packets, how many equations do you have? Do you have enough equations to solve for all of the unknowns? (Think about the answer to the earlier question - does it make sense now why we send as many packets as we do?)

(g) If you have $Q(x)$ and $E(x)$, how does one recover $P(x)$? If you know $P(x)$, how can you recover the original message?

2 Berlekamp-Welch Algorithm

Note 8
Note 9

In this question we will send the message $(m_0, m_1, m_2) = (1, 1, 4)$ of length $n = 3$. We will use an error-correcting code for $k = 1$ general error, doing arithmetic over GF(5).

(a) Construct a polynomial $P(x) \pmod 5$ of degree at most 2, so that

$$P(0) = 1, \qquad P(1) = 1, \qquad P(2) = 4.$$

What is the message $(c_0, c_1, c_2, c_3, c_4)$ that is sent?

(b) Suppose the message is corrupted by changing c_0 to 0. Set up the system of linear equations in the Berlekamp-Welch algorithm to find $Q(x)$ and $E(x)$.

(c) Assume that after solving the equations in part (b) we get $Q(x) = 4x^3 + x^2 + x$ and $E(x) = x$. Show how to recover the original message from Q and E.

3 Berlekamp-Welch Algorithm with Fewer Errors

Note 8
Note 9

In class we derived how the Berlekamp-Welch algorithm can be used to correct k general errors, given $n+2k$ points transmitted. In real life, it is usually difficult to determine the number of errors that will occur. What if we have less than k errors? This is a follow up to the exercise posed in the notes.

Suppose Alice wants to send 1 message to Bob and wants to guard against 1 general error. She decides to encode the message with $P(x) = 4$ (on GF(7)) such that $P(0) = 4$ is the message she want to send. She then sends $P(0), P(1), P(2) = (4,4,4)$ to Bob.

(a) Suppose Bob receives the message $(4,5,4)$. Without performing Gaussian elimination explicitly, find $E(x)$ and $Q(x)$.

(b) Now, suppose there were no general errors and Bob receives the original message $(4,4,4)$. Show that the $Q(x), E(x)$ that you found in part (a) still satisfies $Q(i) = r_i E(i)$ for all $i = 0,1,2$.

(c) Verify that $E(x) = x$, $Q(x) = 4x$ is another possible set of polynomials that satisfies $Q(i) = r_i E(i)$ for all $i = 0,1,2$.

(d) Suppose you're actually trying to decode the received message $(4,4,4)$. Based on what you showed in the previous two parts, what will happen during row reduction when you try to solve for the unknowns?

(e) Prove that in general, no matter what the solution of $Q(x)$ and $E(x)$ are though, the recovered $P(x)$ will always be the same.

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

DIS 5B

1 Countability: True or False

Note 11

(a) The set of all irrational numbers $\mathbb{R}\backslash\mathbb{Q}$ (i.e. real numbers that are not rational) is uncountable.

(b) The set of integers x that solve the equation $3x \equiv 2 \pmod{10}$ is countably infinite.

(c) The set of real solutions for the equation $x + y = 1$ is countable.

For any two functions $f : Y \to Z$ and $g : X \to Y$, let their composition $f \circ g : X \to Z$ be given by $(f \circ g)(x) = f(g(x))$ for all $x \in X$. Determine if the following statements are true or false.

(d) f and g are injective (one-to-one) $\implies$ $f \circ g$ is injective (one-to-one).

(e) f is surjective (onto) $\implies$ $f \circ g$ is surjective (onto).

2 Counting Cartesian Products

Note 11

For two sets A and B, define the cartesian product as $A \times B = \{(a,b) : a \in A, b \in B\}$.

(a) Given two countable sets A and B, prove that $A \times B$ is countable.

(b) Given a finite number of countable sets $A_1, A_2, \ldots, A_n$, prove that

$$A_1 \times A_2 \times \cdots \times A_n$$

is countable.

(c) Consider a countably infinite number of finite sets: $B_1, B_2, \ldots$ for which each set has at least 2 elements. Prove that $B_1 \times B_2 \times \cdots$ is uncountable.

3 Hello World!

Note 12

Determine the computability of the following tasks. If it's not computable, write a reduction or self-reference proof. If it is, write the program.

(a) You want to determine whether a program P on input x prints "Hello World!". Is there a computer program that can perform this task? Justify your answer.

(b) You want to determine whether a program P prints "Hello World!" before running the kth line in the program. Is there a computer program that can perform this task? Justify your answer.

(c) You want to determine whether a program P prints "Hello World!" in the first k steps of its execution. Is there a computer program that can perform this task? Justify your answer.

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 6A**

1 Hello World!

Note 12

Determine the computability of the following tasks. If it's not computable, write a reduction or self-reference proof. If it is, write the program.

(a) You want to determine whether a program P on input x prints "Hello World!". Is there a computer program that can perform this task? Justify your answer.

(b) You want to determine whether a program P prints "Hello World!" before running the kth line in the program. Is there a computer program that can perform this task? Justify your answer.

(c) You want to determine whether a program P prints "Hello World!" in the first k steps of its execution. Is there a computer program that can perform this task? Justify your answer.

2 Code Reachability

Note 12

Consider triplets (M, x, L) where

- `M` is a Java program
- `x` is some input
- `L` is an integer

and the question of: if we execute $M(x)$, do we ever hit line L?

Prove this problem is undecidable.

3 Strings

Note 10

What is the number of strings consisting of:

(a) n ones, and m zeroes?

(b) n_1 A's, n_2 B's and n_3 C's?

(c) $n_1, n_2, \ldots, n_k$ respectively of k different letters?

4 You'll Never Count Alone

Note 10

(a) An anagram of LIVERPOOL is any re-ordering of the letters of LIVERPOOL, i.e., any string made up of the letters L, I, V, E, R, P, O, O, L in any order. For example, IVLERPOOL and POLIVOLRE are anagrams of LIVERPOOL but PIVEOLR and CHELSEA are not. The anagram does not have to be an English word.

How many different anagrams of LIVERPOOL are there?

(b) How many solutions does $y_0 + y_1 + \cdots + y_k = n$ have, if each y must be a non-negative integer?

(c) How many solutions does $y_0 + y_1 + \cdots + y_k = n$ have, if each y must be a positive integer?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 6B**

1 Inclusion and Exclusion

Note 10 What is the total number of positive integers strictly less than 100 that are also coprime to 100?

2 CS70: The Musical

Note 10 Edward, one of the previous head TA's, has been hard at work on his latest project, *CS70: The Musical*. It's now time for him to select a cast, crew, and directing team to help him make his dream a reality.

(a) First, Edward would like to select directors for his musical. He has received applications from $2n$ directors. Use this to provide a combinatorial argument that proves the following identity:

$$\binom{2n}{2} = 2\binom{n}{2} + n^2.$$

(b) Edward would now like to select a crew out of n people. Use this to provide a combinatorial argument that proves the following identity: (this is called Pascal's Identity)

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

(c) There are n actors lined up outside of Edward's office, and they would like a role in the musical (including a lead role). However, he is unsure of how many individuals he would like to cast. Use this to provide a combinatorial argument that proves the following identity:

$$\sum_{k=1}^{n} k \binom{n}{k} = n2^{n-1}$$

(d) Generalizing the previous part, provide a combinatorial argument that proves the following identity:

$$\sum_{k=j}^{n} \binom{n}{k}\binom{k}{j} = 2^{n-j}\binom{n}{j}.$$

3 Farmer's Market

Note 10

Suppose you want k items from the farmer's market. Count how many ways you can do this, assuming:

(a) There are pumpkins and apples at the market.

(b) There are pumpkins, apples, oranges, and pears at the market.

(c) There are n kinds of fruits at the market, and you want to end up with at least two different types of fruit.

4 The Count

Note 10

(a) The Count is trying to choose his new 7-digit phone number. Since he is picky about his numbers, he wants it to have the property that the digits are non-increasing when read from left to right. For example, 9973220 is a valid phone number, but 9876545 is not. How many choices for a new phone number does he have?

(b) Now instead of non-increasing, they must be strictly decreasing. So 9983220 is no longer valid, while 9753210 is valid. How many choices for a new phone number does he have now?

(c) The Count now wants to make a password to secure his phone. His password must be exactly 10 digits long and can only contain the digits 0 and 1. On top of that, he also wants it to contain at least five consecutive 0's. How many possible passwords can he make?

CS 70 Spring 2024 — Discrete Mathematics and Probability Theory — Seshia, Sinclair — DIS 7B

1 Venn Diagram

Note 13

Out of 1,000 computer science students, 400 belong to a club (and may work part time), 500 work part time (and may belong to a club), and 50 belong to a club and work part time.

(a) Suppose we choose a student uniformly at random. Let C be the event that the student belongs to a club and P the event that the student works part time. Draw a picture of the sample space Ω and the events C and P.

(b) What is the probability that the student belongs to a club?

(c) What is the probability that the student works part time?

(d) What is the probability that the student belongs to a club AND works part time?

(e) What is the probability that the student belongs to a club OR works part time?

2 Flippin' Coins

Note 13

Suppose we have an unbiased coin, with outcomes H and T, with probability of heads $\mathbb{P}[H] = 1/2$ and probability of tails also $\mathbb{P}[T] = 1/2$. Suppose we perform an experiment in which we toss the coin 3 times. An outcome of this experiment is (X_1, X_2, X_3), where $X_i \in \{H, T\}$.

(a) What is the *sample space* for our experiment?

(b) Which of the following are examples of *events*? Select all that apply.

- $\{(H,H,T),(H,H),(T)\}$
- $\{(T,H,H),(H,T,H),(H,H,T),(H,H,H)\}$
- $\{(T,T,T)\}$
- $\{(T,T,T),(H,H,H)\}$
- $\{(T,H,T),(H,H,T)\}$

(c) What is the complement of the event $\{(H,H,H),(H,H,T),(H,T,H),(H,T,T),(T,T,T)\}$?

(d) Let A be the event that our outcome has 0 heads. Let B be the event that our outcome has exactly 2 heads. What is $A \cup B$?

(e) What is the probability of the outcome (H,H,T)?

(f) What is the probability of the event that our outcome has exactly two heads?

(g) What is the probability of the event that our outcome has at least one head?

3 Sampling

Note 13

Suppose you have balls numbered $1,\ldots,n$, where n is a positive integer ≥ 2, inside a coffee mug. You pick a ball uniformly at random, look at the number on the ball, replace the ball back into the coffee mug, and pick another ball uniformly at random.

(a) What is the probability that the first ball is 1 and the second ball is 2?

(b) What is the probability that the second ball's number is strictly less than the first ball's number?

(c) What is the probability that the second ball's number is exactly one greater than the first ball's number?

(d) Now, assume that after you looked at the first ball, you did *not* replace the ball in the coffee mug (instead, you threw the ball away), and then you drew a second ball as before. Now, what are the answers to the previous parts?

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

DIS 8A

1 Box of Marbles

Note 14

You are given two boxes: one of them containing 900 red marbles and 100 blue marbles, the other one contains 500 red marbles and 500 blue marbles.

(a) If we pick one of the boxes randomly, and pick a marble what is the probability that it is blue?

(b) If we see that the marble is blue, what is the probability that it is chosen from box 1?

(c) Suppose we pick one marble from box 1 and without looking at its color we put it aside. Then we pick another marble from box 1. What is the probability that the second marble is blue?

2 Poisoned Smarties

Note 14

Supposed there are 3 people who are all owners of their own Smarties factories. Burr Kelly, being the brightest and most innovative of the owners, produces considerably more Smarties than her competitors and has a commanding 50% of the market share. Yousef See, who inherited her riches, lags behind Burr and produces 40% of the world's Smarties. Finally Stan Furd, brings up the rear with a measly 10%. However, a recent string of Smarties related food poisoning has forced the FDA investigate these factories to find the root of the problem. Through her investigations, the inspector found that 2 Smarties out of every 100 at Kelly's factory was poisonous. At See's factory, 5% of Smarties produced were poisonous. And at Furd's factory, the probability a Smarty was poisonous was 0.1.

(a) What is the probability that a randomly selected Smarty will be safe to eat?

(b) If we know that a certain Smarty didn't come from Burr Kelly's factory, what is the probability that this Smarty is poisonous?

(c) Given this information, if a randomly selected Smarty is poisonous, what is the probability it came from Stan Furd's Smarties Factory?

3 Pairwise Independence

Note 14

Recall that the events A_1, A_2, and A_3 are *pairwise independent* if for all $i \neq j$, A_i is independent of A_j. However, pairwise independence is a weaker statement than *mutual independence*, which requires the additional condition that $\mathbb{P}[A_1 \cap A_2 \cap A_3] = \mathbb{P}[A_1]\mathbb{P}[A_2]\mathbb{P}[A_3]$.

Suppose you roll two fair six-sided dice. Let A_1 be the event that the first die lands on 1, let A_2 be the event that the second die lands on 6, and let A_3 be the event that the two dice sum to 7.

(a) Compute $\mathbb{P}[A_1]$, $\mathbb{P}[A_2]$, and $\mathbb{P}[A_3]$.

(b) Are A_1 and A_2 independent?

(c) Are A_2 and A_3 independent?

(d) Are A_1, A_2, and A_3 pairwise independent?

(e) Are A_1, A_2, and A_3 mutually independent?

1 Probability Potpourri

Note 13
Note 14

Provide brief justification for each part.

(a) For two events A and B in any probability space, show that $\mathbb{P}[A \setminus B] \geq \mathbb{P}[A] - \mathbb{P}[B]$.

(b) Suppose $\mathbb{P}[D \mid C] = \mathbb{P}[D \mid \overline{C}]$, where $\overline{C}$ is the complement of C. Prove that D is independent of C.

(c) If A and B are disjoint, does that imply they're independent?

2 Easter Eggs

Note 14

You made the trek to Soda for a Spring Break-themed homework party, and every attendee gets to leave with a party favor. You're given a bag with 20 chocolate eggs and 40 (empty) plastic eggs. You pick 5 eggs (uniformly) without replacement.

(a) What is the probability that the first egg you drew was a chocolate egg?

(b) What is the probability that the second egg you drew was a chocolate egg?

(c) Given that the first egg you drew was an empty plastic one, what is the probability that the fifth egg you drew was also an empty plastic egg?

3 Balls and Bins

Note 14

Suppose you throw n balls into n labeled bins one at a time.

(a) What is the probability that the first bin is empty?

(b) What is the probability that the first k bins are empty?

(c) Let A be the event that at least k bins are empty. Let m be the number of subsets of k bins out of the total n bins. If we assume A_i is the event that the ith set of k bins is empty. Then we can write A as the union of A_i's:

$$A = \bigcup_{i=1}^{m} A_i.$$

Compute m in terms of n and k, and use the union bound to give an upper bound on the probability $\mathbb{P}[A]$.

(d) What is the probability that the second bin is empty given that the first one is empty?

(e) Are the events that “the first bin is empty” and “the first two bins are empty” independent?

(f) Are the events that “the first bin is empty” and “the second bin is empty” independent?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 9A**

1 Head Count

Note 15

Consider a coin with $\mathbb{P}[\text{Heads}] = 2/5$. Suppose you flip the coin 20 times, and define X to be the number of heads.

(a) What is $\mathbb{P}[X = k]$, for some $0 \leq k \leq 20$?

(b) Name the distribution of X and what its parameters are.

(c) What is $\mathbb{P}[X \geq 1]$? Hint: You should be able to do this without a summation.

(d) What is $\mathbb{P}[12 \leq X \leq 14]$?

2 Family Planning

Note 15

Mr. and Mrs. Johnson decide to continue having children until they either have their first girl or until they have three children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let G denote the numbers of girls that the Johnsons have. Let C be the total number of children they have.

(a) Determine the sample space, along with the probability of each sample point.

(b) Compute the joint distribution of G and C. Fill in the table below.

	$C=1$	$C=2$	$C=3$
$G=0$			
$G=1$			

(c) Use the joint distribution to compute the marginal distributions of G and C and confirm that the values are as you'd expect. Fill in the tables below.

$\mathbb{P}[G=0]$	
$\mathbb{P}[G=1]$	

$\mathbb{P}[C=1]$	$\mathbb{P}[C=2]$	$\mathbb{P}[C=3]$

(d) Are G and C independent?

(e) What is the expected number of girls the Johnsons will have? What is the expected number of children that the Johnsons will have?

3 Pullout Balls

Note 15

Suppose you have a bag containing four balls numbered $1, 2, 3, 4$.

(a) You perform the following experiment: pull out a single ball and record its number. What is the expected value of the number that you record?

(b) You repeat the experiment from part (a), except this time you pull out two balls together and record the product of their numbers. What is the expected value of the total that you record?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 9B**

1 Linearity

Note 15

Solve each of the following problems using linearity of expectation. Explain your methods clearly.

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears? (*Hint*: Consider where the sequence "book" can appear in the string.)

2 Head Count II

Note 19

Consider a coin with $\mathbb{P}[\text{Heads}] = 3/4$. Suppose you flip the coin until you see heads for the first time, and define X to be the number of times you flipped the coin.

(a) What is $\mathbb{P}[X = k]$, for some $k \geq 1$?

(b) Name the distribution of X and what its parameters are.

(c) What is $\mathbb{P}[X > k]$, for some $k \geq 0$?

(d) What is $\mathbb{P}[X < k]$, for some $k \geq 1$?

(e) What is $\mathbb{P}[X > k \mid X > m]$, for some $k \geq m \geq 0$? How does this relate to $\mathbb{P}[X > k - m]$?

(f) 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.

3 Shuttles and Taxis at Airport

Note 19

In front of terminal 3 at San Francisco Airport is a pickup area where shuttles and taxis arrive according to a Poisson distribution. The shuttles arrive at a rate $\lambda_1 = 1/20$ (i.e. 1 shuttle per 20 minutes) and the taxis arrive at a rate $\lambda_2 = 1/10$ (i.e. 1 taxi per 10 minutes) starting at 00:00. The shuttles and the taxis arrive independently.

(a) What is the distribution of the following:

(i) The number of taxis that arrive between times 00:00 and 00:20?

(ii) The number of shuttles that arrive between times 00:00 and 00:20?

(iii) The total number of pickup vehicles that arrive between times 00:00 and 00:20?

(b) What is the probability that exactly 1 shuttle and 3 taxis arrive between times 00:00 and 00:20?

(c) Given that exactly 1 pickup vehicle arrived between times 00:00 and 00:20, what is the conditional probability that this vehicle was a taxi?

(d) Suppose you reach the pickup area at 00:20. You learn that you missed 3 taxis and 1 shuttle in those 20 minutes. What is the probability that you need to wait for more than 10 mins until either a shuttle or a taxi arrives?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair DIS 11A

1 Student Life

Note 19

In an attempt to avoid having to do laundry often, Marcus comes up with a system. Every night, he designates one of his shirts as his dirtiest shirt. In the morning, he randomly picks one of his shirts to wear. If he picked the dirtiest one, he puts it in a dirty pile at the end of the day (a shirt in the dirty pile is not used again until it is cleaned).

When Marcus puts his last shirt into the dirty pile, he finally does his laundry, and again designates one of his shirts as his dirtiest shirt (laundry isn't perfect) before going to bed. This process then repeats.

(a) If Marcus has n shirts, what is the expected number of days that transpire between laundry events? Your answer should be a function of n involving no summations.

(b) Say he gets even lazier, and instead of organizing his shirts in his dresser every night, he throws his shirts randomly onto one of n different locations in his room (one shirt per location), designates one of his shirts as his dirtiest shirt, and one location as the dirtiest location.

In the morning, if he happens to pick the dirtiest shirt, *and* the dirtiest shirt was in the dirtiest location, then he puts the shirt into the dirty pile at the end of the day and does not throw any future shirts into that location and also does not consider it as a candidate for future dirtiest locations (it is too dirty).

What is the expected number of days that transpire between laundry events now? Again, your answer should be a function of n involving no summations.

2 Elevator Variance

Note 16

A building has n upper floors numbered $1,2,\dots,n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each person gets off at one of the n upper floors uniformly at random and independently of everyone else. What is the *variance* of the number of floors the elevator *does not* stop at?

3 Covariance

Note 16

(a) We have a bag of 5 red and 5 blue balls. We take two balls uniformly at random from the bag without replacement. Let X_1 and X_2 be indicator random variables for the events of the first and second ball being red, respectively. What is $\text{cov}(X_1, X_2)$? Recall that $\text{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$.

(b) Now, we have two bags A and B, with 5 red and 5 blue balls each. Draw a ball uniformly at random from A, record its color, and then place it in B. Then draw a ball uniformly at random from B and record its color. Let X_1 and X_2 be indicator random variables for the events of the first and second draws being red, respectively. What is $\operatorname{cov}(X_1, X_2)$?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair DIS 11B

1 Probabilistic Bounds

Note 17

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

2 Vegas

Note 17

On the planet Vegas, everyone carries a coin. Many people are honest and carry a fair coin (heads on one side and tails on the other), but a fraction p of them cheat and carry a trick coin with heads on both sides. You want to estimate p with the following experiment: you pick a random sample of n people and ask each one to flip their coin. Assume that each person is independently likely to carry a fair or a trick coin.

(a) Let X be the proportion of coin flips which are heads. Find $\mathbb{E}[X]$.

(b) Given the results of your experiment, how should you estimate p? (*Hint:* Construct an unbiased estimator for p using part (a). Recall that $\hat{p}$ is an unbiased estimator if $\mathbb{E}[\hat{p}] = p$.)

(c) How many people do you need to ask to be 95% sure that your answer is off by at most 0.05?

3 Working with the Law of Large Numbers

Note 17

(a) A fair coin is tossed multiple times and you win a prize if there are more than 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(b) A fair coin is tossed multiple times and you win a prize if there are more than 40% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(c) A fair coin is tossed multiple times and you win a prize if there are between 40% and 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(d) A fair coin is tossed multiple times and you win a prize if there are exactly 50% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

1 Continuous Intro

Note 21

(a) Is

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

a valid density function? Why or why not? Is it a valid CDF? Why or why not?

(b) Calculate the PDF $f_X(x)$, along with $\mathbb{E}[X]$ and $\mathrm{Var}(X)$ if the CDF of X is

$$F_X(x) = \begin{cases} 0, & x \le 0 \\ \frac{x}{\ell}, & 0 \le x \le \ell, \\ 1, & x \ge \ell \end{cases}$$

(c) Suppose X and Y are independent and have densities

$$f_X(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{otherwise}, \end{cases} \qquad f_Y(y) = \begin{cases} 1, & 0 \le y \le 1, \\ 0, & \text{otherwise}. \end{cases}$$

What is their joint distribution? (Hint: for parts (c) and (d), we can use independence in much the same way that we did in discrete probability)

(d) Calculate $\mathbb{E}[XY]$ for the X and Y in part (c).

2 Darts Again

Note 21

Edward and Khalil are playing darts on a circular dartboard.

Edward's throws are uniformly distributed over the entire dartboard, which has a radius of 10 inches. Khalil has good aim (but his throws may land outside of the dartboard); the distance of his throws from the center of the dartboard follows an exponential distribution with parameter $\frac{1}{2}$.

Say that Edward and Khalil both throw one dart at the dartboard. Let X be the distance of Edward's dart from the center, and Y be the distance of Khalil's dart from the center of the dartboard. What is $\mathbb{P}[X < Y]$, the probability that Edward's throw is closer to the center of the board than Khalil's? Leave your answer in terms of an unevaluated integral.

[*Hint:* X is not uniform over $[0,10]$. Solve for the distribution of X by first computing the CDF of X, $\mathbb{P}[X < x]$.]

3 Lunch Meeting

Note 21

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.

(a) Provide a sketch of the joint distribution of the arrival times of Alice and Bob. For which region of the graph will Alice and Bob actually meet?

(b) Based on your sketch, what is the probability that they will actually meet for lunch?

CS 70 Discrete Mathematics and Probability Theory
Spring 2024 Seshia, Sinclair **DIS 12B**

1 Interesting Gaussians

Note 21

(a) If $X \sim N(0, \sigma_X^2)$ and $Y \sim N(0, \sigma_Y^2)$ are independent, then what is $\mathbb{E}\left[(X+Y)^k\right]$ for any *odd* $k \in \mathbb{N}$?

(b) Let $f_{\mu,\sigma}(x)$ be the density of a $N(\mu, \sigma^2)$ random variable, and let X be distributed according to $\alpha f_{\mu_1,\sigma_1}(x) + (1-\alpha) f_{\mu_2,\sigma_2}(x)$ for some $\alpha \in [0,1]$. Compute $\mathbb{E}[X]$ and $\mathrm{Var}(X)$. Is X normally distributed?

2 Binomial Concentration

Note 21

Here, we will prove that the binomial distribution is *concentrated* about its mean as the number of trials tends to ∞. Suppose we have i.i.d. trials, each with a probability of success $1/2$. Let S_n be the number of successes in the first n trials (n is a positive integer).

(a) Compute the mean and variance of S_n.

(b) How should we define Z_n in terms of S_n to ensure that Z_n has mean 0 and variance 1?

(c) What is the distribution of Z_n as $n \to \infty$?

(d) Use the bound $\mathbb{P}[Z > z] \leq (\sqrt{2\pi}z)^{-1}\,\mathrm{e}^{-z^2/2}$ when Z is a standard normal in order to approximately bound $\mathbb{P}[S_n/n > 1/2 + \delta]$, where $\delta > 0$.

3 Erasures, Bounds, and Probabilities

Note 21

Alice is sending 1000 bits to Bob. The probability that a bit gets erased is p, and the erasure of each bit is independent of the others.

Alice is using a scheme that can tolerate up to one-fifth of the bits being erased. That is, as long as Bob receives at least 801 of the 1000 bits correctly, he can decode Alice's message.

In other words, Bob becomes unable to decode Alice's message only if 200 or more bits are erased. We call this a "communication breakdown", and we want the probability of a communication breakdown to be at most 10^{-6}.

(a) Use Chebyshev's inequality to upper bound p such that the probability of a communications breakdown is at most 10^{-6}.

(b) As the CLT would suggest, approximate the fraction of erasures by a Gaussian random variable (with suitable mean and variance). Use this to find an approximate bound for p such that the probability of a communications breakdown is at most 10^{-6}.

You may use that $\Phi^{-1}(1-10^{-6}) \approx 4.753$.

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

Discussion 13A

1 Markov Chain Basics

Note 22

A Markov chain is a sequence of random variables X_n, $n = 0, 1, 2, \ldots$. Here is one interpretation of a Markov chain: X_n is the state of a particle at time n. At each time step, the particle can jump to another state. Formally, a Markov chain satisfies the Markov property:

$$\mathbb{P}[X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0] = \mathbb{P}[X_{n+1} = j \mid X_n = i], \tag{1}$$

for all n, and for all sequences of states $i_0, \ldots, i_{n-1}, i, j$. In other words, the Markov chain does not have any memory; the transition probability only depends on the current state, and not the history of states that have been visited in the past.

(a) In lecture, we learned that we can specify Markov chains by providing three ingredients: $\mathscr{X}$, P, and π_0. What do these represent, and what properties must they satisfy?

(b) If we specify $\mathscr{X}$, P, and π_0, we are implicitly defining a sequence of random variables X_n, $n = 0, 1, 2, \ldots$, that satisfies (1). Explain why this is true.

(c) Calculate $\mathbb{P}[X_1 = j]$ in terms of π_0 and P. Then, express your answer in matrix notation. What is the formula for $\mathbb{P}[X_n = j]$ in matrix form?

2 Can it be a Markov Chain?

Note 22

(a) A fly flies in a straight line in unit-length increments. Each second it moves to the left with probability 0.3, right with probability 0.3, and stays put with probability 0.4. There are two spiders at positions 1 and m and if the fly lands in either of those positions it is captured. Given that the fly starts between positions 1 and m, model this process as a Markov Chain.

(b) Take the same scenario as in the previous part with $m = 4$. Let $Y_n = 0$ if at time n the fly is in position 1 or 2 and let $Y_n = 1$ if at time n the fly is in position 3 or 4. Is the process Y_n a Markov chain?

3 Allen's Umbrella Setup

Note 22

Every morning, Allen walks from his home to Soda, and every evening, Allen walks from Soda to his home. Suppose that Allen has two umbrellas in his possession, but he sometimes leaves his umbrellas behind. Specifically, before leaving from his home or Soda, he checks the weather. If it is raining outside, he will bring exactly one umbrella (that is, if there is an umbrella where he currently is). If it is not raining outside, he will forget to bring his umbrella. Assume that the probability of rain is p.

(a) Model this as a Markov chain. What is $\mathscr{X}$? Write down the transition matrix.

(b) What is the transition matrix after 2 trips? n trips? Determine if the distribution of X_n converges to the invariant distribution, and compute the invariant distribution.

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

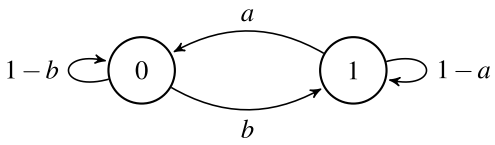
Discussion 13B

1 Markov Chain Terminology

Note 22

In this question, we will walk you through terms related to Markov chains.

1. (Irreducibility) A Markov chain is irreducible if, starting from any state i, the chain can transition to any other state j, possibly in multiple steps.
2. (Periodicity) $d(i) := \gcd\{n > 0 \mid P^n(i,i) = \mathbb{P}[X_n = i \mid X_0 = i] > 0\}$, $i \in \mathscr{X}$. If $d(i) = 1\ \forall i \in \mathscr{X}$, then the Markov chain is aperiodic; otherwise it is periodic.
3. (Matrix Representation) Define the transition probability matrix P by filling entry (i, j) with probability $P(i, j)$.
4. (Invariance) A distribution π is invariant for the transition probability matrix P if it satisfies the following balance equations: $\pi = \pi P$.



(a) For what values of a and b is the above Markov chain irreducible? Reducible?

(b) For $a = 1$, $b = 1$, prove that the above Markov chain is periodic.

(c) For $0 < a < 1$, $0 < b < 1$, prove that the above Markov chain is aperiodic.

(d) Construct a transition probability matrix using the above Markov chain.

(e) Write down the balance equations for this Markov chain and solve them. Assume that the Markov chain is irreducible.

2 Skipping Stones

Note 22

We consider a simple Markov chain model for skipping stones on a river, but with a twist: instead of trying to make the stone travel as far as possible, you want the stone to hit a target. Let the set of states be $\mathscr{X} = \{1,2,3,4,5\}$. State 3 represents the target, while states 4 and 5 indicate that you have overshot your target. Assume that from states 1 and 2, the stone is equally likely to skip forward one, two, or three steps forward. If the stone starts from state 1, compute the probability of reaching our target before overshooting, i.e. the probability of $\{3\}$ before $\{4,5\}$.

3 Consecutive Flips

Note 22

Suppose you are flipping a fair coin (one Head and one Tail) until you get the same side 3 times (Heads, Heads, Heads) or (Tails, Tails, Tails) in a row.

(a) Construct an Markov chain that describes the situation with a start state and end state.

(b) Given that you have flipped a (Tails, Heads) so far, what is the expected number of flips to see the same side three times?

(c) What is the expected number of flips to see the same side three times, beginning at the start state?

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

1 Short Tree Proofs

Note 5

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

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

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

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

2 Secret Sharing Practice

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

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

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

3 Counting Subsets

Note 11

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

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

(b) Prove or disprove: S is countable.

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

Prove that F is countably infinite.

4 Strings

Note 10

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

5 Mario's Coins

Note 14

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

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

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

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

6 Sum of Poisson Variables

Note 19

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

Hint: Recall the binomial theorem.

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

7 Balls in Bins

Note 15

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

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

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

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

8 Inequality Practice

Note 17

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

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

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

9 Exponential Distributions: Lightbulbs

Note 21

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

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

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

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

10 Continuous Probability Continued

Note 17

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)?

11 Three Tails

Note 22

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

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