CS 70 Discrete Mathematics and Probability Theory
Spring 2025 Rao

Final Solutions

PRINT Your Name: Oski Bear

SIGN Your Name: $\mathscr{OSKI}$

Do not turn this page until your instructor tells you to do so.

1. **Pledge.**

Berkeley Honor Code: As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others.

In particular, I acknowledge that:

- I alone am taking this exam. Other than with the course staff, I will not have any verbal, written, or electronic communication about the exam with anyone else while I am taking the exam or while others are taking the exam.
- I will not refer to any books, notes, or online sources of information while taking the exam, other than what the instructor has allowed.
- I will not take screenshots, photos, or otherwise make copies of exam questions to share with others.

SIGN Your Name: ______________________________

2. **Warmup.**

(1 point) May 16th is a special day: not only is it the day of the Spring 2025 CS70 final exam, but it is also which TA's birthday? (Hint: It may help to read through the entire exam!)

Answer: Jay: this is mentioned in the last question of the exam!

3. **Propositional Logic**

1. Let P, Q and R be propositions.

(a) $(P \implies \neg Q) \equiv (\neg P \vee \_\_\_\_)$ (Fill in the blank to make the statement always true.)
Answer: $\neg Q$. We have that $P \implies Q \equiv \neg P \vee Q$, so in this case, we have that $P \implies \neg Q \equiv \neg P \vee \neg Q$.

(b) $P \wedge (Q \vee R) \equiv (P \vee Q) \wedge (P \vee R)$.
Answer: False. Suppose P is False, then the LHS must be False, but if Q and R are both true, then the RHS is True.

(c) $P \wedge (\neg P \vee Q) \wedge (\neg P \vee \neg Q)$
Answer: Always false.
If P is false, then the whole proposition is immediately false. Elsewise, if P is true, then both $(\neg P \vee Q)$ and $(\neg P \vee \neg Q)$ can be simplified to (False $\vee Q) \equiv Q$ and (False $\vee \neg Q) \equiv \neg Q$ respectively. We can see that these cannot both be true simultaneously, and thus their conjunction must be False.

2. Let $P(n)$ and $Q(n)$ are propositions over $n \in \mathbb{N}$.

(a) Suppose the following is true:

$$(\forall n \in \mathbb{N})(P(n) \implies Q(n+1)) \wedge (\forall n \in \mathbb{N})(Q(n) \implies P(n+1)) \wedge P(0)$$

Determine whether the following statements are true or false.

i. $(\forall n \in \mathbb{N})(2 \mid n \implies P(n))$
Answer: True. $P(n) \implies Q(n+1) \implies P(n+2)$ and $P(0)$, or $P(0)$ and $P(2i) \implies P(2(i+1))$, which by the principle of induction is true for all i.

ii. $(\forall n \in \mathbb{N})(\neg Q(n+1) \implies \neg P(n))$
Answer: True. This is the contrapositive of $P(n) \implies Q(n+1)$.

iii. $Q(0) \implies (\forall n \in \mathbb{N})(Q(n))$.
Answer: True, the original propositions make all even $P(n)$ true and all odd $Q(n)$ true, and having $Q(0)$ be true as well fills the rest of the inductive gaps.

(b) $\neg(\forall n \in \mathbb{N})(P(n) \vee Q(n)) \implies (\exists n \in \mathbb{N})(\neg P(n))$.
Answer: True. There must be a value of n where both $P(n)$ and $Q(n)$ are false.

4. Proofs

1. (8 points) Suppose p and q are distinct primes larger than 2.
Prove that if $e \equiv d^{-1} \pmod{(p-1)(q-1)}$, then $x^{ed} \equiv x \pmod{pq}$.
(You are welcome to use the Chinese Remainder Theorem and/or Fermat's Little Theorem, but simply asserting this is true from the fact that RSA works properly will earn 0 points.)
Answer: Since $ed \equiv 1 \pmod{(p-1)(q-1)}$, we know that $ed = k(p-1)(q-1)+1$.
This means that we'd like to show that

$$\begin{aligned} x^{ed} \equiv x^{k(p-1)(q-1)+1} &\equiv x \pmod{pq} \\ x^{k(p-1)(q-1)+1} - x &\equiv 0 \pmod{pq} \\ x(x^{k(p-1)(q-1)} - 1) &\equiv 0 \pmod{pq} \end{aligned}$$

Let us consider this equivalence under mod p and under mod q separately.
Under mod p, we want to look at $x(x^{k(p-1)(q-1)}-1) \pmod{p}$. Note that there are only two ways in which this expression can be zero: either $x \equiv 0 \pmod{p}$, or $x^{k(p-1)(q-1)} - 1 \equiv 0 \pmod{p}$. Hence, if $x \equiv 0 \pmod{p}$, we immediately find that this expression is equivalent to 0; as such, we'll focus on the case where $x \not\equiv 0 \pmod{p}$.
Here, we'd like to show that $x^{k(p-1)(q-1)} \equiv 1 \pmod{p}$. Since $x \not\equiv 0 \pmod{p}$, we have by FLT that

$$\begin{aligned} x^{k(p-1)(q-1)} &= (x^{p-1})^{k(q-1)} \pmod{p} \\ &\equiv 1^{k(q-1)} \equiv 1 \pmod{p} \end{aligned}$$

which shows that $x^{k(p-1)(q-1)} - 1 \equiv 0 \pmod{p}$ as desired.
The same argument applies under mod q; the expression is equivalent to 0 immediately if $x \equiv 0 \pmod{q}$, and otherwise for $x \not\equiv 0$, we have by FLT that

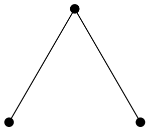
$$\begin{aligned} x^{k(p-1)(q-1)} &= (x^{q-1})^{k(p-1)} \pmod{q} \\ &\equiv 1^{k(p-1)} \equiv 1 \pmod{q} \end{aligned}$$

This means that we have the system

$$\begin{cases} x(x^{k(p-1)(q-1)}-1) \equiv 0 \pmod{p} \\ x(x^{k(p-1)(q-1)}-1) \equiv 0 \pmod{q} \end{cases} \implies x(x^{k(p-1)(q-1)}) \equiv 0 \pmod{pq}$$

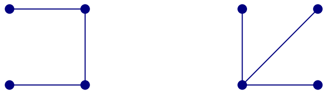
by CRT, completing the proof.

2. A graph $H = (V_H, E_H)$ is a *subgraph* of another graph $G = (V_G, E_G)$ if there exists an injective function $h : V_H \to V_G$ mapping vertices in H to vertices in G, such that each edge $(u,v) \in E_H$ corresponds to an edge $(h(u), h(v)) \in E_G$.
There is only one tree on three vertices; the length two path:



(a) (4 points) Draw all of the 4-vertex trees.

Answer: There are only two 4-vertex trees:



(b) Every three vertex tree is a subgraph of K_3.

Answer: True. Every three vertex graph is a subgraph of K_3.

(c) (10 points) Prove by induction on k that any simple graph G with minimum degree d contains every k-vertex tree as a subgraph, for each $1 \le k \le d$.

Answer: The 1-vertex tree is a single vertex, and is a subgraph of all (non-empty) graphs G.

We assume as our inductive hypothesis that for a fixed $k < d$, all trees with at most k vertices is a subgraph of G.

For our inductive step, we show that the $(k+1)$-vertex tree T is also a subgraph of G. We know that T must contain a leaf v, with its neighbor u; we can remove leaf v from T, forming a k-vertex tree T'. By our inductive hypothesis, T' is a subgraph of G.

Here, we can observe that $h(u)$ must have at least one free neighboring vertex; there are at most $k-1 \le d$ neighbors of u in T', so at most $k-1 \le d$ neighbors of $h(u)$ have already been assigned a corresponding vertex in T. This leaves at least one neighbor w of $h(u)$ unassigned.

As such, we can define $h(v) = w$, completing the mapping h; this means that the $(k+1)$-vertex tree T is a subgraph of G, as desired.

5. Stable Matching

1. In a stable matching instance with n jobs and n candidates, if all jobs have the same preference list, how many days does it take for the job propose and reject algorithm to terminate?

   Answer: n. Everyone asks the first candidate, and all but one ask the second and the rest. Also accepting $n-1$ as the termination date may be the day when everyone gets one proposal or the day before.

2. If all jobs have the same preference list, there is only one stable matching.

   Answer: True. The first candidate can choose their favorite job. If there was a matching where the candidate and job are not matching then they would be a rogue couple. Thus, neither entity in this pair can be a rogue couple. And by induction, there is only one stable pairing in the instance on the remaining $n-1$ jobs and candidates.

3. If all candidates have the same preference list, there is only one stable matching.

   Answer: True. The argument is symmetric. Note the algorithm is not relevant to the argument.

4. Consider a stable matching instance with n jobs and n candidates, where the job and candidate preference lists are uniformly random permutations.

   (a) What is the expected number of candidates who get exactly i proposals on the first day? (Hint: the binomial distribution might be useful.)

   Answer: $n\binom{n}{i}(\frac{1}{n})^i(1-\frac{1}{n})^{n-i}$. For each candidate c, the probability that it is first in job j's list is $1/n$. There are n independent jobs that could've placed candidate c at the top of their preference list, so the distribution of proposals that candidate c gets is $\text{Bin}(n, p = 1/n)$ and thus the probability that candidate c gets exactly i proposals is $\binom{n}{i}(\frac{1}{n})^i(1-\frac{1}{n})^{n-i}$.

Denoting X as the total number of candidates that receive exactly i proposals and X_l as an indicator for whether the lth candidate received exactly i proposals, we get $\mathbb{E}[X] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] = n\mathbb{E}[X_l] = n\binom{n}{i}(\frac{1}{n})^i(1-\frac{1}{n})^{n-i}$.

(b) For $n = 2$, recall that there are only two possible matchings. What is the probability that both are stable?

Answer: $\frac{1}{8}$. For both to be stable the jobs must prefer different candidates. This is because, if J_1 and J_2 both prefer C_1, then whichever job C_1 prefers must be paired with C_1 (and then there would only be one possible stable matching), else they are a rogue couple. So the job preference lists must be either $J_1 : C_1 > C_2, J_2 : C_2 > C_1$ or $J_1 : C_2 > C_1, J_2 : C_1 > C_2$.

Now, notice that if J_1 prefers C_1, then C_1 must not prefer J_1, else they must be paired together (and then there would only be one possible stable matching). Thus, there are only two possible sets of preference lists that result in both two stable matchings, out of a total of 16 possible sets.

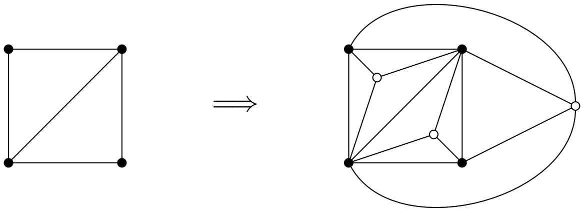
6. Graphs

1. Given an n vertex planar graph where every face has exactly four edges, then the graph has exactly _______ edges. (Your answer may possibly be in terms of n.)

   Answer: $2n-4$. In this case the number of face-edge adjacencies is both $4f$ and $2e$ or $f = e/2$, and plugging into euler's, $n + e/2 = e + 2$, or $e = 2n-4$.

2. Let G be a simple planar graph; in the following, we will consider a planar drawing of G. We define the *face-vertex graph* $F = (V_F, E_F)$ of G to be the graph that contains all vertices in G alongside an additional vertex for each face in G. Every edge in G also appears in F, but we additionally define edges incident to face vertices such that $(f, v) \in E_F$ if the face f (which is a cycle in G) contains the vertex v.

   For example, we have the following drawing of a graph G and its face-vertex graph F (face vertices are in white).



   In the following, let F be the face vertex graph for a planar drawing of a graph G with v vertices, f faces, and e edges. You may assume that F is planar.

   (a) How many vertices are in the face-vertex graph of G?

   Answer: $v + f$. There is a vertex for each original vertex and for each face.

   (b) How many edges are in the face-vertex graph of G? (Hint: What is the length of each face?)

   Answer: $3(v + f) - 6$. One can draw the face-vertex graph in a planar manner, and the length of every cycle is 3. Such a graph has exactly $3N - 6$ edges where N is the number of vertices. In this case the number of vertices is $v + f$.

3. A tree has one connected component. Removing an edge results in a graph with ________ connected component(s).

   Answer: 2. There are no cycles, so the endpoints of the removed edge will be in separate components. No other part of the graph is modified, so we end up with exactly 2 connected components.

4. To make K_{2n} bipartite, one must remove at least ________ edges. (Give a tight bound; your answer may possibly be in terms of n.)

   Answer: $\binom{2n}{2} - n^2 = 2\binom{n}{2}$. K_{2n} contains $\binom{2n}{2}$ edges and $K_{n,n}$ contains n^2 edges. Any other bipartite graph contains fewer edges.

   Alternatively, this can be accomplished by arbitrarily partitioning the K_{2n} into two sets of size n and removing the edges between vertices in the same set, for which there are $\binom{n}{2}$ edges in each set, resulting in $2 \times \binom{n}{2}$ total edges removed. Note that this is equivalent to the previous answer, since $\binom{2n}{2} - n^2 = \frac{2n(2n-1)}{2} - n^2 = n(2n-1) = n^2 = n(n-1) = 2 \times \binom{n}{2}$

5. The least number of colors to *vertex color* a tree with max degree d is ________. (Your answer may possibly be in terms of d.)

   Answer: 2. There is always a vertex of degree 1 in the tree. So remove it, recursively color the tree with two colors. One can then color the "removed" vertex with one of the two colors as it is adjacent to only one vertex.

6. The least number of colors to *edge color* a degree d tree is ________. (Your answer may possibly be in terms of d.)

   Answer: d. Begin with an arbitrary vertex, its neighboring edges can be colored with d colors. Then each neighbor has at most 1 colored edge and $d-1$ other edges which can be colored with the remaining $d-1$ colors. Since there are no cycles, this can be applied recursively to each neighbor.

   This is also a tight bound, since the star graph with one vertex of degree d and all other vertices of degree 1 requires d colors for an edge coloring.

7. The maximum number of edges in an n vertex graph where every cycle has length at least n is ________. (Your answer may possibly be in terms of n.)

   Answer: n. If it has a cycle of length n, adding any other edge creates a smaller cycle.

8. The length of an Eulerian tour in an Eulerian graph with n vertices and m edges is ________. (Your answer may possibly be in terms of n and/or m.)

   Answer: m. It must contain all the edges once.

9. Consider removing a simple cycle of length k from a connected graph with m edges and n vertices. The resulting graph has c connected components.

   (a) What is the minimum value of c? (Your answer may possibly be in terms of k and/or n. Do not use m.)

   Answer: 1. It may not disconnect the graph at all.

   (b) What is the maximum value of c? (Your answer may possibly be in terms of k and/or n. Do not use m.)

   Answer: k. Each vertex could end up in a separate component.

7. Modular Arithmetic

In the following parts, when working under arithmetic modulo N, your answers should be given in the range $\{0, 1, \ldots, N-1\}$.

1. What is $2^{36} \pmod 7$?

   Answer: $1 \pmod 7$. By FLT, $2^6 \equiv 1 \pmod 7$ so $(2^6)^6 \equiv 1 \pmod 7$.

2. What is $2^{36} \pmod{35}$?

Answer: $1 \pmod{35}$. $2^{36} \equiv 1 \pmod{7}$ as above. $2^{36} \equiv 1 \pmod{5}$ since $2^4 \equiv 1 \pmod{5}$ (by FLT). This means that we have the system

$$\begin{cases} 2^{36} \equiv 1 \pmod{5} \\ 2^{36} \equiv 1 \pmod{7} \end{cases} \implies 2^{36} \equiv 1 \pmod{35}$$

by CRT.

3. If $x \equiv 0 \pmod{d}$, then $bx - kd \equiv 0 \pmod{d}$ for any values of b and k.

   Answer: True. Here, we have

$$bx - kd \equiv bx \equiv b \cdot 0 = 0 \pmod{d}.$$

4. If $x \equiv 1 \pmod{d}$, then $bx - kd \equiv 1 \pmod{d}$ for any values of b and k.

   Answer: False. Here, we have

$$bx - kd \equiv bx \equiv b \pmod{d}$$

   which is not always equal to $1 \pmod{d}$.

5. Find $\gcd(385, 70)$.

   Answer: 35. Two steps of the iterative GCD shows this. Also $70 = 35 \times 2$ and $385 = 11 \times 35$. Since $2 \nmid 385$, 35 is the greatest common divisor.

6. Let $\gcd(a, m) = d$, and $2a \equiv b \pmod{m}$.

   (a) Then $(2 + \_\_\_\_)a \equiv b \pmod{m}$. (Your answer may possibly be in terms of a, d, and/or m, and must not be $0 \pmod{m}$.)

   Answer: m/d. This is equivalent to solving for x in $(2+x)a \equiv b \pmod{m}$. Here, we have

$$\begin{aligned} (2+x)a &\equiv b \pmod{m} \\ 2a + ax &\equiv b \pmod{m} \\ b + ax &\equiv b \pmod{m} \\ ax &\equiv 0 \pmod{m} \\ ax &= km \qquad (k \in \mathbb{Z}) \end{aligned}$$

   Since $\gcd(a, m) = d$, we know that a must be a multiple of d, and m must be a multiple of d. This means that $\frac{a}{d}$ and $\frac{m}{d}$ are both integers; dividing by d on both sides now, we have

$$\begin{aligned} ax &= km \\ \frac{a}{d}x &= k\frac{m}{d} \\ \frac{a}{d}x &\equiv 0 \pmod{\frac{m}{d}} \end{aligned}$$

   Here, any $x \equiv 0 \pmod{\frac{m}{d}}$ is a solution. In particular, $x = \frac{m}{d}$ is a solution (we cannot choose $x = 0$).

   (b) $2 \mid b$.

   Answer: False. For $a = 6, b = 3, m = 9$ we have $2 \times 6 = 3 \pmod{9}$ and $2 \nmid 3 = b$.

   (c) $d \mid b$.

   Answer: True. $a = id$ and $m = jd$ for $i, j \in Z$, and $2a \equiv b \pmod{m}$ implies that there is an $\ell \in Z$ such that $2a + \ell m = 2id + \ell jd = b$, implies $d|b$.

(d) The number of solutions to $ax \equiv b \pmod{m}$ is ________ if there is at least 1 solution.
Answer: d.
Similar to part (a), if $ax \equiv b \pmod{m}$, then we have $ax = b + km$ for some $k \in \mathbb{Z}$. Since $\gcd(a,m) = d$, both a and m are multiples of d; this forces b to also be a multiple of d if there is to be any solution to the equation.
Since all quantities in the equality are divisible by d, we can divide through by d to get

$$\frac{a}{d}x = \frac{b}{d} + k\frac{m}{d} \implies \frac{a}{d}x \equiv \frac{b}{d} \pmod{\frac{m}{d}}$$

Since now $\gcd(\frac{a}{d}, \frac{m}{d}) = 1$, there is a unique solution for x under mod $\frac{m}{d}$. In particular, this means that $\{x, x + \frac{m}{d}, x + 2\frac{m}{d}, \ldots, x + (d-1)\frac{m}{d}\}$ are all solutions under mod m, giving d total solutions.

7. A perfect square modulo m is a value x such that there exists an a such that $a^2 \equiv x \pmod{m}$.

(a) How many perfect squares are there under modulo 3?
Answer: 2. Listing out all the squares, we have

$$\begin{aligned} 0^2 &\equiv 0 \pmod{3} \\ 1^2 &\equiv 1 \pmod{3} \\ 2^2 &\equiv 4 \equiv 1 \pmod{3} \end{aligned}$$

This means that the only perfect squares are 0 and 1 under modulo 3.

(b) If $p > 2$ is prime, how many perfect squares are there under modulo p?
Answer: $1 + \frac{p-1}{2}$.
In order for a value a to be a perfect square, there must exist some x such that $x^2 \equiv a \pmod{p}$. In other words, there must exist a root of $x^2 - a \pmod{p}$.
The polynomial $x^2 - a \pmod{p}$ has at most 2 roots under modulo p. In fact, these roots must be of the form $x \equiv r \pmod{p}$ and $x \equiv -r \pmod{p}$.
The only case when the two roots coincide is when

$$r \equiv -r \pmod{p} \implies 2r \equiv 0 \pmod{p} \implies r \equiv 0 \pmod{p},$$

so $r = 0$ is the only case where we have a double root. In all other cases, we have pairs of distinct roots $(r, -r)$ that correspond to the same perfect square r^2.
This means that we can pair up all of the values from $\{1, 2, \ldots, p-1\}$, creating exactly $\frac{p-1}{2}$ distinct perfect squares. Adding in 0 as a perfect square as well, this gives us $1 + \frac{p-1}{2}$ distinct perfect squares under modulo p.

(c) How many perfect squares are there under modulo 15? (It might be easier to do the next problem.)
Answer: 6.
Checking all possible squares, we have

$$\begin{aligned} 0^2 &= 0 \pmod{15} \\ 1^2 &= 1 \pmod{15} \\ 2^2 &= 4 \pmod{15} \\ 3^3 &= 9 \pmod{15} \\ 4^2 &= 16 \equiv 1 \pmod{15} \\ 5^2 &= 25 = 10 \pmod{15} \\ 6^2 &= 36 \equiv 6 \pmod{15} \\ 7^2 &= 49 \equiv 4 \pmod{15} \end{aligned}$$

$$\begin{aligned} 8^2 &= 64 \equiv 4 \pmod{15} \\ 9^2 &= 81 \equiv 6 \pmod{15} \\ 10^2 &= 100 \equiv 10 \pmod{15} \\ 11^2 &= 121 \equiv 1 \pmod{15} \\ 12^2 &= 144 \equiv 9 \pmod{15} \\ 13^2 &= 169 \equiv 4 \pmod{15} \\ 14^2 &= 196 \equiv 1 \pmod{15} \end{aligned}$$

There are 6 perfect squares here: $\{0,1,4,6,9,10\}$.
Alternatively, the next subpart suggests that the number of perfect squares mod 15 is equal to the product of the number of perfect squares under mod 3 and mod 5. We know from part (a) that there are 2 perfect squares mod 3, and we can observe that there are 3 perfect squares mod 5 (i.e. $\{0,1,4\}$). Their product is 6, as desired.

(d) If $m = pq$ for distinct primes $p, q > 2$, how many perfect squares are there under modulo m? (Hint: CRT. You may also use the function $P(r)$ as the number of perfect squares under modulo r for a prime r, with appropriate arguments in your solution.)
Answer: $(1+\frac{p-1}{2})(1+\frac{q-1}{2}) = P(p)P(q)$.
From the previous parts, we have that the number of perfect squares modulo p is $P(p) = 1+\frac{p-1}{2}$ and the number of perfect squares modulo q is $P(q) = 1+\frac{q-1}{2}$.
We will show that for any value y, y is a perfect square modulo pq if and only if it is a perfect square modulo p and modulo q.
Forward direction: If $y \equiv x^2 \pmod{pq}$ for some integer x, then $y \equiv x^2 \pmod{p}$ and $y \equiv x^2 \pmod{q}$, so y is a perfect square modulo both p and q.
Reverse direction: Suppose $y \equiv a^2 \pmod{p}$ and $y \equiv b^2 \pmod{q}$ for some integers a, b.
Since p and q are distinct primes, the Chinese Remainder Theorem guarantees that there exists an integer x such that $x \equiv a \pmod{p}$ and $x \equiv b \pmod{q}$. Then $x^2 \equiv a^2 \equiv y \pmod{p}$ and $x^2 \equiv b^2 \equiv y \pmod{q}$. Therefore, by the Chinese Remainder Theorem again, $x^2 \equiv y \pmod{pq}$, so y is a perfect square modulo pq.
Thus, picking y to be a perfect square modulo pq is equivalent to picking $a^2 \pmod{p}$, $b^2 \pmod{q}$, for which there are $P(p)P(q)$ ways to do so.

8. Polynomials.

1. (1 point each) Your friend Faith is performing Lagrange interpolation through three distinct points in the field GF(5). She has already found the interpolating polynomial $P(x) = x^2 + 4x + 1$ and you know that one of her *delta polynomials* is $\Delta(x) = x^2 + 3x + 2$.
   Determine the three original points she must have used, in the form (x, y).

   Answer: Evaluate $\Delta(x) = x^2 + 3x + 2$ at $x = \{0,1,2,3,4\}$ in GF(5) or factor it into $(x+2)(x+1)$ to see that the zeros are $x = 3$ and $x = 4$. There is one other point $x = i$ that we must find, for which this delta polynomial corresponds to. Recall that by definition, $\Delta_i(x) = \frac{(x-3)(x-4)}{(i-3)(i-4)}$ so we have that $(i-3)(i-4) \equiv 1 \pmod{5}$. The only point in GF(5) that satisfies this is $i = 1$.
   Now that we have the x-values, we can substitute them back into $P(x) = x^2 + 4x + 1$ to get the y-values, resulting in $(1,1)$, $(3,2)$ and $(4,3)$.
2. Any degree exactly 2 polynomial is a bijection under modulo p for a prime p. (Hint: think about perfect squares.)
   Answer: False. For x^2, both a and $-a$ are in the pre-image of a^2.
3. Any degree exactly 1 polynomial is a bijection under modulo p for a prime p.
   Answer: True. For any line $ax + b \pmod{p}$, for an image $y \equiv ax + b \pmod{p}$ we have $x \equiv a^{-1}(y-b) \pmod{p}$ which is unique as a has a unique multiplicative inverse.
4. Recall that two polynomials $P(x)$ and $Q(x)$ intersect at a if $P(a) = Q(a)$.

(a) Find the x-value where $P(x) = 2x+3 \pmod 5$ and $Q(x) = 3x+2 \pmod 5$ intersect. (That is, find a value a where $P(a) \equiv Q(a) \pmod 5$.)
Answer: $x = 1$. Solve $2x+3 \equiv 3x+2 \pmod 5$ for x.

(b) If $P(x)$ and $Q(x)$ are distinct polynomials, with degrees d_p and d_q respectively, what is the maximum possible number of intersections under modulo p for a prime $p > \max(d_p, d_q)$? (In other words, at most how many distinct values of a are there such that $P(a) \equiv Q(a) \pmod p$? Give a tight bound.)
Answer: $\max(d_p, d_q)$. This is the number of roots of $P(x) - Q(x)$. Since $P(x) - Q(x)$ has degree $\max(d_p, d_q)$, and the maximum number of roots of any polynomial is its degree, there is a maximum of $\max(d_p, d_q)$ intersections.

(c) Working modulo a prime $p > 3$, suppose we fix a polynomial $P(x)$ of degree 3. How many degree 3 polynomials $Q(x)$ are there such that $Q(1) \equiv P(1) \pmod p$ and $Q(2) \equiv P(2) \pmod p$?
Answer: p^2. Since $Q(x)$ is of degree 3, 4 points uniquely define Q. Two of these points are fixed by $P(1)$ and $P(2)$, so we have two additional points to select; there are p^2 possibilities for these two remaining points.

5. Recall that the Berlekamp–Welch algorithm encodes a message using a polynomial $P(x)$, and decodes using the error polynomial $E(x)$.

(a) Suppose you want to send a message of length n, and would like to tolerate k corruptions. What is the degree of $P(x)$?
Answer: $n-1$. We encode each message as a point $(i, P(i))$, and n points are sufficient to determine a degree $n-1$ polynomial.

(b) Suppose you want to send a message of length n, and would like to tolerate k corruptions. How many points $(i, P(i))$ would you be sending?
Answer: $n+2k$. n packets to recover the polynomial. k packets will have no information. And one needs k more for various reasons having to do with locating the errors. It is sufficient because any polynomial consistent with $n+k$ received points must be $P(x)$.

(c) (5 points) Suppose you receive the packets (i, r_i), with at most k corruptions (i.e. $r_i \neq P(i)$ at most k times.)
Argue that $E(i)(P(i) - r_i) = 0$ at all of the points that are sent. (You should use properties of the error polynomial $E(x)$ which corresponds to the errors in r_i.)
Answer: $E(i) = 0$ at the k points where there is an error. And when there is no error, we have $P(i) = r_i$ or that $P(i) - r_i = 0$. Thus, $E(i)(P(i) - r_i)$ is always zero as the first term is 0 when there is an error and the second term is 0 when there isn't.

9. Counting

Throughout this question, you may leave your answers unsimplified (i.e. you can leave binomial coefficients, factorials, exponents, etc. as is), but you should not use any summation or product notation (i.e. you may not use $\sum$ or $\prod$).

1. How many ways can the letters in BAA be arranged?
Answer: $3 = \frac{3!}{2}$. 3! counts the possible arrangements, but we divide by 2 to account for the fact that the A's are identical.

2. How many ways can the digits in 126 be arranged?
Answer: $6 = 3!$. 3 ways to choose the first one, 2 ways to choose the second one, and 1 way to choose the last.

3. Suppose there are n teddy bears and k children, for $n > k$. How many ways are there to assign teddy bears to the children, such that every child has a single distinct teddy bear? The bears and children are both distinguishable.

   Answer: $\frac{n!}{(n-k)!}$. There are n ways to assign to the first one, $n-1$ to the second, etc.

4. How many simple bipartite graphs $G = (L, R, E)$ are there with m total edges, such that $|L| = |R| = n$ and $E \subseteq L \times R$? Here, assume that vertices are labeled.

   Answer: $\binom{|L|^2}{m} = \binom{|R|^2}{m}$. Choose m out of the $|L|^2$ edges.

5. How many functions are there under modulo p, for a prime p? (A function under modulo p maps $f : \{0, \ldots, p-1\} \to \{0, \ldots, p-1\}$.)

   Answer: You can think of the function as a p length string from a size p alphabet, for which there is a count of p^p.

6. How many bijective functions are there under modulo p, for a prime p?

   Answer: Thinking as a string again, the string must be a permutation, for which there are $p!$ counts.

7. How many ways are there to make 3 teams of 4 players out of 15 players in total? (The teams are distinguishable, i.e., assume the teams are named as Teams 1, 2 and 3.)

   Answer: $\binom{15}{4}\binom{11}{4}\binom{7}{4}$. The first term is the number of ways to form team 1, the second term is the number of ways to form team 2 from the remaining 11 players, and the third term corresponds to choosing the third team from the remaining 7 players.

10. Combinatorial Proof.

1. (8 points) Give a combinatorial proof that $\sum_{i=0}^{n} \binom{n}{i}\binom{n}{n-i} = \binom{2n}{n}$.

   Answer: The right hand side is choosing n items out of $2n$. On the left hand side, one arbitrarily splits the elements into two groups of n elements and choosing i from one group and $n-i$ from the other for each possible i. Any subset of size n is counted exactly once as any subset contains exactly i elements from one half and $n-i$ from the other.

2. (4 points) Argue the previous part implies that $\sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}$.

   Answer: Replace $\binom{n}{n-i}$ by $\binom{n}{i}$ as each corresponds to the number of subsets of size i either by choosing a subset or by choosing what's left out of a subset.

11. Countability

1. The cardinality of the set of all subsets of the natural numbers is the same as the cardinality of the set of real numbers in the interval $[0, 1]$.

   Answer: True. One can create a bijection by noting that a real number in $[0, 1]$ can be represented in binary (with a decimal point) and is thus the same as an an infinite string of 0's and 1. Then a subset S of the natural numbers is specified by the string where 0 in the ith digit means $i \notin S$ and a 1 signifies that the $i \in S$.

2. If $A_1, A_2, \ldots, A_n$ are countable sets, than the $\mathscr{A} = \{(a_1, \ldots, a_n) : a_i \in A_i\}$ is countable.

   Answer: True. One can enumerate the set by enumerating the set $\mathscr{A}$ in order of the sum of the indices of the elements in their original sets.

3. Let A, B, and C be sets, with $A = \mathbb{Z} \times \mathbb{Z}$. Suppose there exists a surjective (onto) function $f : A \to C$, and an injective (one-to-one) function $g : C \to B$.

(a) A is countable.

Answer: True. There is a bijection from $\mathbb{Z} \times \mathbb{Z}$ to the natural numbers, by traversing the 2D grid of integers in a spiral.

(b) The cardinality of A is strictly smaller than the cardinality of B.

Answer: False.

The surjective function implies that $|A| \geq |C|$, and the injective function implies that $|C| \leq |B|$. These two inequalities by themselves do not say anything about the relative cardinalities of A and B. In particular, it's possible that $C = \mathbb{N}$, and $B = \mathbb{R}$, or that $C = B = \mathbb{N}$.

4. Give a bijection from the reals in $(0,1]$ to the reals in $[1,\infty)$.

$f(x) =$

Answer: $f(x) = 1/x$. It is its own inverse and maps numbers less than 1 to number greater than 1 and 1 to 1.

12. Computability

1. There exists a program which given another program P, an integer k, and an input x, determines whether the program $P(x)$ halts in at most $|x|^k$ steps.

   Answer: True. One can just run P (in an interpreter) on x for up to $|x|^k$ steps. Answer yes if it halts.

2. Suppose you have a program `HaltsOnEmpty`, which given an program P, determines whether P halts on the empty string.

   Fill in the following blanks to write a program `Halt` that takes a program P and an input x and determines whether P halts on input x. You may only use variables defined in the template.

```
def Halt(P, x):
    def inner(y):
        _______(1)_______

    return HaltsOnEmpty(______(2)______)
```

(1) (2)

Answer:

```
def Halt(P, x):
    def inner(y):
        P(x)

    return HaltsOnEmpty(inner)
```

13. Basic Probability.

Consider a sample space Ω with the probability function $\mathbb{P} : \Omega \to (0,1]$, and events $A, B, C \subseteq \Omega$ all with non-zero size.

All of your answers below may only be expressions involving $\mathbb{P}[A]$, $\mathbb{P}[B]$, $\mathbb{P}[C]$, and/or some real numbers.

1. Give the lowest upper bound on $\mathbb{P}[A \cup B]$.

   Answer: $\mathbb{P}[A] + \mathbb{P}[B]$. This is the union bound and the best possible upper bound without using $\mathbb{P}[A \cap B]$.

2. Give the highest lower bound on $\mathbb{P}[A \cup B]$.

   Answer: $\max(\mathbb{P}[A], \mathbb{P}[B])$, since the union at least contains all the elements of A, and the elements of B.

3. If $A \subseteq B$, give the highest lower bound on $\mathbb{P}[A \mid B]$.

   Answer: $\frac{\mathbb{P}[A]}{\mathbb{P}[B]}$. $A \cap B = A$.

4. If $A \subseteq B$, give the highest lower bound on $\mathbb{P}[B \mid A]$.

   Answer: 1. Any sample point in A is in B.

5. Suppose $\mathbb{P}[A \cap B] = \alpha$ and $\mathbb{P}[B \cap C] = \alpha$.

   Your answers below may now additionally involve α. (Hint: Think about the sets for various α's.)

   (a) Give the lowest upper bound on $\mathbb{P}[A \cap C]$.

   Answer: $\min(\mathbb{P}[A], \mathbb{P}[C])$. We have $A \cap C \subseteq A$, so by monotonicity $\mathbb{P}[A \cap C] \leq \mathbb{P}[A]$. Similarly, $\mathbb{P}[A \cap C] \leq \mathbb{P}[C]$.

   (b) Give the highest lower bound on $\mathbb{P}[A \cap C]$.

   Answer: $\max(0, 2\alpha - \mathbb{P}[B])$.

   (c) (3 points) Justify your answer for the previous question.

   Answer: We have by total probability and nonnegativity of probability

   $$\mathbb{P}[A \cap C] = \mathbb{P}[A \cap C \cap B] + \mathbb{P}[A \cap C \cap B^c] \geq \mathbb{P}[A \cap C \cap B]$$

   Also,

   $$\mathbb{P}[A \cap C \cap B] = \mathbb{P}[(A \cap B) \cap (C \cap B)].$$

   Now using inclusion-exclusion,

   $$\mathbb{P}[(A \cap B) \cap (C \cap B)] = \mathbb{P}[A \cap B] + \mathbb{P}[C \cap B] - \mathbb{P}[(A \cap B) \cup (C \cap B)].$$

   We are given $\mathbb{P}[A \cap B] = \mathbb{P}[C \cap B] = \alpha$. Also, $(A \cap B) \cup (C \cap B) = (A \cup C) \cap B$ by distribution, and $\mathbb{P}[(A \cup C) \cap B] \leq \mathbb{P}[B]$ by monotonicity. Substituting this all in, we get

   $$\mathbb{P}[A \cap C] \geq 2\alpha - \mathbb{P}[B].$$

   However, for a small enough α, this quantity can go below zero. Since probabilities are always nonnegative, we can take zero as a fixed lower bound, giving us a final answer of $\max(0, 2\alpha - \mathbb{P}[B])$.

14. Four or six?

Suppose you have a four sided die and a six sided die. You choose one die uniformly at random and roll it twice. Let A be the event that the first roll is 1, and let B be the event that the second roll is a 1.

1. What is $\mathbb{P}[A \mid B]$?

   Answer: 13/60.

   $$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = \frac{\frac{1}{2} \times \frac{1}{16} + \frac{1}{2} \times \frac{1}{36}}{\frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{6}} = \frac{\frac{1}{2} \times \frac{13}{144}}{\frac{1}{2} \times \frac{10}{24}} = \frac{13}{144} \times \frac{24}{10} = \frac{13}{60}.$$

2. Let C be the event that the four sided die is chosen. What is $\mathbb{P}[C \mid A]$?

   Answer: $3/5$.

$$\mathbb{P}[C \mid A] = \frac{\mathbb{P}[A \cap C]}{\mathbb{P}[A]} = \frac{\mathbb{P}[C]\mathbb{P}[A \mid C]}{\mathbb{P}[C]\mathbb{P}[A \mid C] + \mathbb{P}[\overline{C}]\mathbb{P}[A \mid \overline{C}]} = \frac{\frac{1}{2} \times \frac{1}{4}}{\frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{6}} = \frac{3}{5}.$$

15. Bernoulli Random Variables

1. Consider two independent Bernoulli random variables X and Y with $\mathbb{E}[X] = \mathbb{E}[Y] = 1/2$. What is $\mathbb{P}[X = Y]$?

   Answer: $1/2$. If $\mathbb{E}[X] = \mathbb{E}[Y] = \frac{1}{2}$, then X, Y are independently distributed as Bernoulli$(\frac{1}{2})$, with equal probability of taking on either 0 or 1.

   This means that we have $\mathbb{P}[X = Y] = \mathbb{P}[X = 0 \cap Y = 0] + \mathbb{P}[X = 1 \cap Y = 1] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$.

2. Consider two Bernoulli random variables X and Y with $\mathbb{E}[X] = \mathbb{E}[Y] = 1/2$, such that $\text{corr}(X,Y) = 0.5$.

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

   Answer: $3/8$.

   Since $X, Y \sim$ Bernoulli$(\frac{1}{2})$ (here, X and Y are dependent RVs!), we know that $\text{Var}(X) = \text{Var}(Y) = \frac{1}{4}$, and

$$\begin{aligned} \text{corr}(X,Y) &= \frac{\text{cov}(X,Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}} \\ \frac{1}{2} &= \frac{\text{cov}(X,Y)}{\sqrt{\frac{1}{4} \cdot \frac{1}{4}}} \\ \text{cov}(X,Y) &= \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \end{aligned}$$

   Further, since $\text{cov}(X,Y) = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[XY]$, we have

$$\begin{aligned} \frac{1}{8} &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ \mathbb{E}[XY] &= \frac{1}{8} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{8} \end{aligned}$$

   (b) What is $\mathbb{P}[X = Y]$?

   Answer: $3/4$.

   We know from part (a) that $\mathbb{E}[XY] = \mathbb{P}[X = 1 \cap Y = 1] = \frac{3}{8}$.

   Next, we want to compute $\mathbb{P}[X = 1 \cap Y = 0]$. Notice that we have

$$\begin{aligned} \mathbb{P}[X = 1] &= \mathbb{P}[X = 1 \cap Y = 0] + \mathbb{P}[X = 1 \cap Y = 1] \\ \frac{1}{2} &= \mathbb{P}[X = 1 \cap Y = 0] + \frac{3}{8} \\ \mathbb{P}[X = 1 \cap Y = 0] &= \frac{1}{2} - \frac{3}{8} = \frac{1}{8} \end{aligned}$$

Similarly, we have

$$\begin{aligned}\mathbb{P}[Y=1] &= \mathbb{P}[X=0\cap Y=1]+\mathbb{P}[X=1\cap Y=1]\\ \frac{1}{2} &= \mathbb{P}[X=0\cap Y=1]+\frac{3}{8}\\ \mathbb{P}[X=0\cap Y=1] &= \frac{1}{2}-\frac{3}{8}=\frac{1}{8}\end{aligned}$$

The remaining probability must comprise $\mathbb{P}[X=0\cap Y=0]$, so we have

$$\begin{aligned}\mathbb{P}[X=0\cap Y=0] &= 1-(\mathbb{P}[X=1\cap Y=1]+\mathbb{P}[X=0\cap Y=1]+\mathbb{P}[X=1\cap Y=0])\\ &= 1-\left(\frac{3}{8}+\frac{1}{8}+\frac{1}{8}\right)\\ &= 1-\frac{5}{8}=\frac{3}{8}\end{aligned}$$

As such,

$$\mathbb{P}[X=Y]=\mathbb{P}[X=0\cap Y=0]+\mathbb{P}[X=1\cap Y=1]=\frac{3}{8}+\frac{3}{8}=\frac{3}{4}$$

16. Confidence Interval.

Consider the process of sampling n people who tested for flu last year to determine the fraction p of the population that would test positive. To set this up, let X_i be the indicator random variable that person i in the sample tested positive for the flu for $i\in\{1,\dots,n\}$.

1. What is $\mathbb{E}[X_i]$ in terms of p?

   Answer: p. $\mathbb{E}[X_i]$ is just the probability of person i getting the flu.

2. What is the tightest possible upper bound on $\mathrm{Var}(X_i)$, independent of the value of p?

   Answer: $1/4$. The variance of an indicator random variable with expected value p is $p(1-p)$. This is maximized by $p=1/2$.

3. Using Chebyshev's inequality, give a 95% confidence interval for p, given that 50 people in your sample of 100 people tested positive for having the flu.

   Answer: $[1/2-1/\sqrt{20},1/2+1/\sqrt{20}]\approx[.28,.72]$

   The proportion of people who test positive is $X=\frac{1}{100}\sum_{i=1}^{100}X_i$. Since each X_i are identical and independent, we have

   $$\mathrm{Var}(X)=\mathrm{Var}\left(\frac{1}{100}\sum_{i=1}^{100}X_i\right)=\frac{1}{100^2}\sum_{i=1}^{100}\mathrm{Var}(X_i)=\frac{1}{100}\mathrm{Var}(X_1).$$

   We do not know the real probability p that a given person tests positive, so the best bound that we can get is $\mathrm{Var}(X_1)=p(1-p)\leq\frac{1}{4}$.

   This means that we can bound $\mathrm{Var}(X)=\frac{1}{100}\mathrm{Var}(X_1)=\leq\frac{1}{400}$.

   By Chebyshev's inequality, we have

   $$\mathbb{P}[|X-\mathbb{E}[X]|\geq c]\leq\frac{\mathrm{Var}(X)}{c^2}\leq\frac{1}{400c^2},$$

which we want to be at most $0.05 = \frac{1}{20}$. To do this, we can set $c = \frac{1}{\sqrt{20}}$, producing the 95% confidence interval

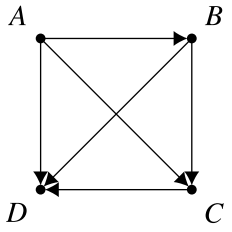
$$p \in \left[\frac{1}{2} - c, \frac{1}{2} + c\right] = \left[\frac{1}{2} - \frac{1}{\sqrt{20}}, \frac{1}{2} + \frac{1}{\sqrt{20}}\right],$$

where the center $\frac{1}{2}$ of the confidence interval comes from our sample of 50 positive cases out of 100.

17. Hamiltonian Paths.

Consider a tournament, i.e. a complete directed graph on n vertices. A Hamiltonian path on a tournament is a sequence of directed edges that visits every vertex exactly once, with each edge pointing from one vertex to the next in the sequence.

For example, the following tournament contains only 1 Hamiltonian path $A \to B \to C \to D$:



1. (8 points) Prove that every tournament contains a Hamiltonian path.

Answer: We proceed with strong induction on the number of vertices.

- Base case $n = 2$: The single edge will bring you from the first vertex to the last vertex, completing the tour.
- Inductive hypothesis: Suppose the statement is true for all tournaments with at most k vertices.
- Inductive step: Consider any tournament with $k+1$ vertices. Pick any vertex v, and define G_{in} to be the set of vertices with edges pointing to v, and G_{out} to be the set of vertices pointing from v. Since $|G_{in}| \leq k, |G_{out}| \leq k$, by the induction hypothesis, there is a Hamiltonian path on G_{in} and G_{out}. Concatenating these paths proves our claim.

Alternatively, we can prove this statement by contradiction. Suppose there exists a tournament that does not contain a Hamiltonian path; in this tournament, suppose we find a longest path in the graph, traversing $v_1 \to v_2 \to \cdots \to v_k$. Since it is not Hamiltonian, there exists some vertex u not in the path. Our goal now is to argue that there exists a longer path in the tournament that includes this new vertex u.

First, let us focus on the first and last vertices in the path. If there is an edge $u \to v_1$, then we can immediately extend the path to be longer, forming a contradiction. This means that there must be an edge $v_1 \to u$. Similarly, if there is an edge $v_k \to u$, then we can immediately extend the path to be longer, forming a contradiction. This means that there must be an edge $u \to v_k$.

Notice that now, we know that there is an edge *away* from v_1 to u, and an edge *toward* v_k from u. Looking at adjacent vertices v_i, v_{i+1} along the path, we claim that there must be a location where there is an edge *away* from v_i to u, and an edge *toward* v_{i+1} from u—this is because the direction of the edges must flip at some point when going from v_1 to v_k, since the edges at the endpoints are in opposite directions.

$$\begin{tikzcd}
v_1 \arrow[r] \arrow[drrrr, bend right] & v_2 \arrow[r] & \cdots \arrow[r] & v_i \arrow[r] \arrow[dr] & v_{i+1} \arrow[r] & \cdots \arrow[r] & v_{k-1} \arrow[r] & v_k \\
 & & \cdots & & u \arrow[u] \arrow[urrr, bend right] & \cdots & &
\end{tikzcd}$$

However, this means that we can insert u between vertices v_i and v_{i+1} along the path, so that we form a longer path $v_1 \to \cdots \to v_i \to u \to v_{i+1} \to \cdots \to v_k$, which gives us our contradiction.

2. Consider the following way of choosing a random tournament T on n vertices: independently for each (unordered) pair of distinct vertices $i, j \in \{1, \ldots, n\}$, flip a coin and include the edge $i \to j$ in the graph if the outcome is heads, and the edge $j \to i$ if tails.

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

   Answer: $2^{\binom{n}{2}}$. There are two possible choices for who wins in each of the $\binom{n}{2}$ pairs.

   (b) What is the expected number of Hamiltonian paths? (Hint: use indicator random variables that indicates whether a permutation forms a Hamiltonian path.)

   Answer: $\frac{n!}{2^{n-1}}$. Let X be the number of Hamiltonian paths in an outcome, and X_i be an indicator random variable for the ith permutation being Hamiltonian. The $\mathbb{P}[X_i = 1] = \frac{1}{2^{n-1}}$ and there are $n!$ permutations and the result follows by linearity of expectation.

18. Linear Regression.

1. For random variables X and Y, the linear regression line of Y given X goes through the origin if and only if $\mathbb{E}[Y] =$ ________.

   Answer: $\mathbb{E}[Y] = \frac{\text{cov}(X,Y)}{\text{Var}(X)}\mathbb{E}[X]$. This follows by plugging in $X = 0, Y = 0$ into the LLSE formula.

2. For a random variable X, $\mathbb{E}[X^2] =$ ________. (Give an answer in terms of $\text{Var}(X)$ and/or $\mathbb{E}[X]$.)

   Answer: $\text{Var}(X) + \mathbb{E}[X]^2$. This follows from the variance definition.

3. For random variables X and Y, we have $\text{cov}(X,Y) = \mathbb{E}[XY]$ if $\mathbb{E}[X] = \mathbb{E}[Y] =$ ________.

   Answer: 0. We have that $\text{cov}(XY) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$. Since $\mathbb{E}[X] = \mathbb{E}[Y]$, we need both to be 0.

4. For independent random variables X and Y, what is the best linear estimator for Y given X, i.e., $\hat{Y}(X)$? (Your answer should be fully simplified.)

   Answer: $\mathbb{E}[Y]$. If X and Y were not independent, we would have that

   $$\hat{Y}(X) = \mathbb{E}[Y] + \frac{\text{cov}(X,Y)}{\text{Var}(X)}(X - \mathbb{E}[X]).$$

   However, since X and Y are independent, we know that $\text{cov}(X,Y) = 0$, meaning the entire second term cancels out, leaving us with just $\hat{Y}(X) = \mathbb{E}[Y]$.

19. Distributions

1. Given independent random variables $X, Y \sim \text{Bin}(n, p)$, what is the distribution of $X + Y$?

   Answer: $\text{Bin}(2n, p)$. It's flipping $2n$ coins with probability p.

2. Given independent random variables $X, Y \sim \text{Geom}(p)$, what is the distribution of $\min(X,Y)$?

   Answer: $\text{Geom}(2p - p^2) = \text{Geom}(1 - (1-p)^2)$. This the length of time it takes for the process of flipping two coins in each step and stopping when one gets a heads in either. At each step, the probability of getting at least one heads is the complement of the probability of getting no heads; this means that $\mathbb{P}[\text{at least one heads}] = 1 - \mathbb{P}[\text{no heads}] = 1 - (1-p)^2 = 2p - p^2$.

3. What is $\lim\limits_{n\to\infty} \binom{n}{i} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i}$?

   Answer: $e^{-\lambda}\frac{\lambda^i}{i!}$. This expression is the PMF of $\text{Bin}(n, \frac{\lambda}{n})$. This distribution converges to $\text{Poisson}(\lambda)$ as $n \to \infty$, so the limit converges to $e^{-\lambda}\frac{\lambda^i}{i!}$.

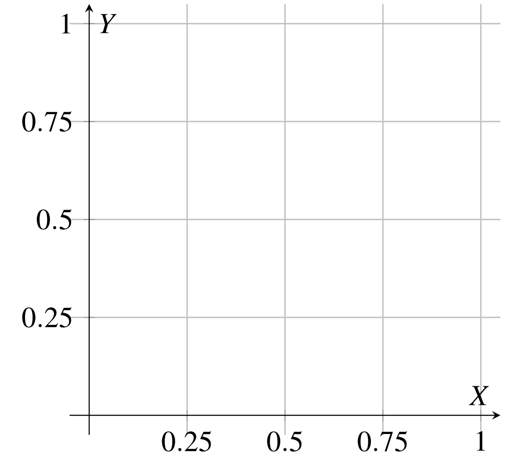
4. Suppose $X \sim \text{Uniform}(0,1]$ and $Y \sim \text{Uniform}[0,X]$.

(a) What is the marginal density function for X? (2 points for the expression, 1 point for bounds)

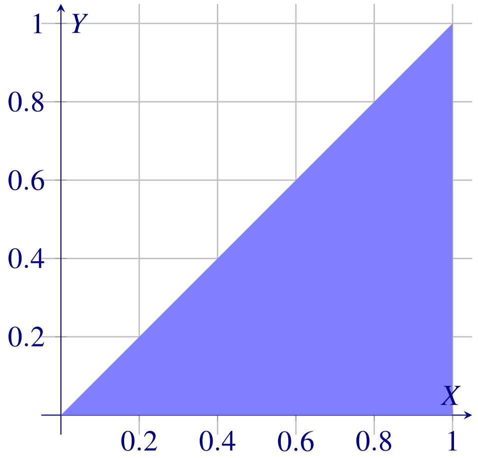
$$f_X(x) = \begin{cases} & \text{if} \\ \\ 0 & \text{otherwise} \end{cases}$$

Answer: $f_X(x) = \begin{cases} 1 & x \in (0,1] \\ 0 & \text{otherwise} \end{cases}$. We choose X first, so $X \sim \text{Uniform}(0,1]$.

(b) (3 points) Shade the region where the density is non-zero.



Answer:



(c) What is the conditional density function $f_{Y|X}(x,y)$ on the shaded region from (b)?

$f_{Y|X}(x,y) =$

Answer: $f_{Y|X}(x,y) = \begin{cases} \frac{1}{x} & x \in [0,1], y \in [0,x] \\ 0 & \text{otherwise} \end{cases}$

This follows from the density being uniform for y given x.

(d) What is the marginal density function for $f_Y(y)$? (2 points for the expression, 1 point for bounds)

$$f_Y(y) = \begin{cases} & \text{if} \\ & \\ 0 & \text{otherwise} \end{cases}$$

Answer: $f(x) = \begin{cases} -\ln(y) & y \in [0,1] \\ 0 & \text{otherwise} \end{cases}$.

We have

$$\int_0^1 f_{Y|X}(x,y) f_X(x)\,\mathrm{d}x = \int_y^1 \frac{1}{x}\,\mathrm{d}x = -\ln(y),$$

for $y \in [0,1]$, and 0 otherwise.

20. Continuous Probability.

1. Let X be a continuous RV with cumulative distribution function (CDF) $F(x)$, and probability density function (PDF) $f(x)$.

   (a) Express $\mathbb{P}[X \in [a,b]]$ in terms of the CDF $F(x)$ for X.
   Answer: $F(b) - F(a)$. The $\mathbb{P}[X \leq b] - \mathbb{P}[X \leq a]$.

   (b) Express $\mathbb{P}[X \in [a,b]]$ in terms of the PDF $f(x)$ for X.
   Answer: $\int_a^b f(x)\,\mathrm{d}x$. This is from the definition of PDF.

   (c) What is $\mathbb{P}[X \leq b \mid X \geq a]$ in terms of the CDF $F(x)$ for X? (Think carefully about the events.)
   Answer: 0 if $a > b$, and otherwise $\frac{F(b)-F(a)}{1-F(a)}$

2. For $X \sim \text{Uniform}[a,b]$.

   (a) What is $\mathbb{P}[X \leq t \mid X \geq s]$ for $a < s < t < b$?
   Answer: $\frac{t-s}{b-s}$. Given that $X \geq s$, the value of X is now uniform on the interval $[s,b]$. The probability that $X \leq t$ under this condition is thus $\frac{t-s}{b-s}$.

   (b) $\text{Uniform}[a,b]$ is memoryless.
   Answer: False. $\mathbb{P}[X \leq \delta = t-s] = \frac{\delta}{b-a} \neq \frac{t-s}{b-s}$.

3. Suppose m real numbers are chosen uniformly and independently at random from $[0,1]$. Let X be the smallest one of these numbers.

   (a) What is the PDF of X? (2 points for expression, 1 point for bounds)

$$f_X(x) = \begin{cases} & \text{if} \\ & \\ 0 & \text{otherwise} \end{cases}$$

Answer: $f_X(x) = \begin{cases} m(1-x)^{m-1} & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$

To find the PDF, we'll find the CDF first; in particular, we'll look at the complement to the CDF. Here, $\mathbb{P}[X > x]$ denotes the probability that the smallest is larger than x; this means that *all* of the chosen numbers must be larger than x. This occurs with probability $(1-x)^m$, making the CDF $\mathbb{P}[X < x] = 1 - (1-x)^m$.

Taking the derivative to find the PDF, we have $f_X(x) = m(1-x)^{m-1}$ when $x \in [0,1]$.
Alternatively, consider the CDF and PDF of one of the uniform random variables U. Then, we claim that $f_X(x) = \binom{m}{1} f_U(x)(1-F_U(x))^{m-1}$. We can derive this combinatorially by first choosing 1 out of the m uniforms to be the smallest (hence the $\binom{m}{1}$), then that value needs to occur (hence $f_U(x)$), and then all other $m-1$ values need to be greater than x (hence $(1-F_U(x))^{m-1}$. Evaluating the expression gives the same answer as before.

(b) What is $\mathbb{E}[X]$?
Answer: $\frac{1}{m+1}$. One can view the problem as cutting a length one necklace (or a numberline from 0 to 1) into $m+1$ pieces, where each of the m real numbers is an endpoint of the necklace (so when combined with the 0 and 1 endpoints, produces $m+1$ pieces). Since the sum of the length of the $m+1$ pieces is 1, and each length has the same expectation by symmetry, then linearity of expectation gives us $\frac{1}{m+1}$. So the first piece, whose endpoint is exactly the smallest of the m uniforms, has expected length $\frac{1}{m+1}$.
Alternatively, by the tail sum for positive random variables, we have

$$\mathbb{E}[X] = \int_0^1 \mathbb{P}[X \geq x]\, \mathrm{d}x = \int_0^1 (1-x)^m\, \mathrm{d}x = \left[-\frac{1}{m+1}(1-x)^{m+1}\right]\Bigg|_0^1 = \frac{1}{m+1}.$$

21. Markov Chain.

1. Consider the Markov Chain on the elements of $\{0,\dots,29\} \pmod{30}$ where there are transitions with equal probability from state $i \pmod{30}$ to state $i+3 \pmod{30}$ and state $i+5 \pmod{30}$.

   (a) Is the chain irreducible?
   Answer: Yes, since you can get to any other state mod 5 with the $i+5$ transition, and within mod 5, you can get to any other state using the $i+3$ transition since 3 and 5 are coprime.

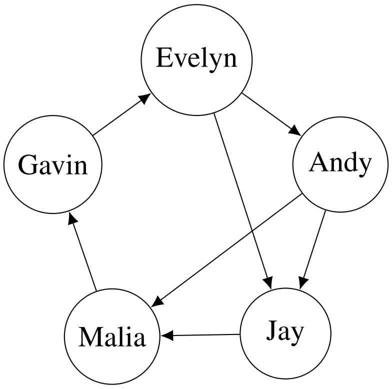
   (b) What is the period of state $0 \pmod{30}$?
   Answer: 2. Notice that with the transitions $i \to i+3$ and $i \to i+5$, we can only go between numbers of opposite parity (i.e. even to odd, and odd to even). This means that any cycle must be of an even length.

   (c) The uniform distribution is invariant.
   Answer: True. Every vertex has two outgoing and two incoming arcs with the same transition probabilities.

2. There is an unexpected pepper shortage at La Val's Pizza! Five friends (Andy, Gavin, Evelyn, Jay, and Malia) must share a single pepper shaker. At time 0, the shaker is in Evelyn's hands, and its location is thereafter recorded at discrete one-minute intervals. During each minute, the current holder passes it to one of the other four friends, according to fixed transition probabilities. The shaker's movement is modeled as a Markov chain with state space

$$S = \{\text{Andy}, \text{Gavin}, \text{Evelyn}, \text{Jay}, \text{Malia}\}.$$

The transitions are depicted below, where each person chooses uniformly over their outgoing arrows.



(a) Compute the expected *hitting time* of Andy, i.e. the expected number of minutes until the shaker first reaches Andy.

Answer: We have the following hitting time equations, where each variable denotes hitting time from the state labelled the first letter of the TA's name, e.g.., $J = \mathbb{E}[\text{time to reach Andy} \mid \text{at Jay}]$.

$$\begin{aligned} J &= 1 + M \\ M &= 1 + G \\ G &= 1 + E \\ E &= 1 + \frac{1}{2}J + \frac{1}{2}A \\ A &= 0 \end{aligned}$$

We eliminate the variable G, M and A (no offense to Gavin, Malia, or Andy) and see that

$$\begin{aligned} J &= 3 + E \\ E &= 1 + \frac{1}{2}J \end{aligned}$$

Substituting J (bye bye Jay) yields $E = 1 + \frac{1}{2}(3 + E)$ and isolating E (hope you are not lonely Evelyn), one obtains $E = 5$.

Alternatively, one can observe that starting from Evelyn, the shaker reaches Andy in one step with probability $\frac{1}{2}$, or with probability $\frac{1}{2}$ one takes 4 steps to get back to Evelyn.

This looks geometric—for each failure, we incur a penalty of 4 steps, and each success adds just 1 step. The expected value for a geometric random variable is the total number of trials that have occurred before getting the first success—for a Geom($\frac{1}{2}$) RV, its expected value is 2. One of these steps must have been the success (we always see exactly one success), so we incur a cost of 1 step for the success. The remaining step must be a failure, so we incur an additional cost of 4 steps for the failure. In total, this allows us to arrive at an expected value of 5 steps before we reach Andy.

(b) It's Jay's birthday today (May 16), and he is worried that Gavin will use all the pepper. Find the probability that the shaker reaches Jay before it next reaches Gavin.

Answer: $\frac{3}{4}$,

We have the following hitting time equations, where each variable denotes hitting time from the state labelled the first letter of the TA's name, e.g., $A = \mathbb{P}[\text{reach Jay before Gavin} \mid \text{at Andy}]$.

$$\begin{aligned} J &= 1 \\ G &= 0 \\ M &= G \\ E &= \frac{1}{2}J + \frac{1}{2}A \\ A &= \frac{1}{2}M + \frac{1}{2}J \end{aligned}$$

Zero and One, my love is won. So, bye bye Gavin and Malia (G and M) who disappoint Jay, and its all done at Jay (J), so we get:

$$\begin{aligned} E &= \frac{1}{2} + \frac{1}{2}A \\ A &= \frac{1}{2} \end{aligned}$$

So from Evelyn, Jay is happy $3/4$ of the time, though Gavin is very nice and I expect Jay worries too much.