CS 70 Fall 2025 — Final, December 19th, 2025

Solutions last updated: 12/20/25

Print Your Name: ______________________

Print Your Student ID: ______________________

Print Student name to your left: ______________________

Print Student name to your right: ______________________

You have three hours. There are 16 questions of varying credit. (154 points total)

Question:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
Points:	4	19	9	10	7	4	18	7	13	3	7	12	17	5	13	6	154

For questions with **circular bubbles**, select only one choice (there is only one correct answer).

- ○ Unselected option (completely unfilled)
- ◎ Don't do this (it will be graded as incorrect)
- ● Only one selected option (completely filled)

For questions with **square boxes**, you may select one or more choices (select all that apply).

- ■ You can select
- ■ multiple squares
- ☑ Don't do this (it will be graded as incorrect)

- There will be no clarifications. We will correct any mistakes post-exam in as fair a manner as possible. Please just answer the question as best you can and move on even if you feel it is a mistake.
- The questions vary in difficulty. In particular, the exam is not in the order of difficulty and quite accessible short answer and proof questions are late in the exam. No points will be given for a blank answer, and there will be no negative points on the exam. **So do really scan over the exam.**
- You may, without proof, use theorems and lemmas that were proven in the notes and/or in lecture, unless otherwise stated. That is, if we ask you to prove a statement, prove it from basic definitions, e.g., "$d \mid x$ means $x = kd$ for some integer k" is a definition.
- You may consult only two double sided sheets of notes on both sides. Apart from that, you may not look at books, notes, etc. Calculators, phones, computers, and other electronic devices are NOT permitted.
- Anything you write outside the answer boxes or you ~~cross out~~ will not be graded. If you write multiple answers or your answer is ambiguous, we will grade the **worst** interpretation.

Pledge

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, large language models, 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: ____

This page intentionally left (mostly) blank

The exam begins on the next page.

Q1 *Illogical Propositions* **(4 points)**

Q1.1 (2 points) $((P \land Q) \Rightarrow R) \equiv (P \Rightarrow (Q \Rightarrow R))$

● Always True ○ Sometimes False

Solution:

P	**Q**	**R**	**P ∧ Q**	**(P ∧ Q) ⇒ R**	**Q ⇒ R**	**P ⇒ (Q ⇒ R)**	**(P ∧ Q ⇒ R) ≡ (P ⇒ Q ⇒ R)**
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	T
T	F	T	F	T	T	T	T
T	F	F	F	T	T	T	T
F	T	T	F	T	T	T	T
F	T	F	F	T	F	T	T
F	F	T	F	T	T	T	T
F	F	F	F	T	T	T	T

Q1.2 (2 points) $\forall x(P(x) \Rightarrow Q(x)) \equiv (\forall x \neg P(x)) \lor (\forall x Q(x))$

○ Always True ● Sometimes False

Solution: A counter-example for this would be:

$P(1) = F$

$P(2) = T$

$Q(1) = F$

$Q(2) = T$

For both 1 and 2, $P(X) \Rightarrow Q(x)$. However, $(\forall x \neg P(x)) \lor (\forall x Q(x))$ would be false because $P(2) = T$ and $Q(1) = F$.

Q2 *Divisors and Modular Arithmetic* (19 points)

Q2.1 (5 points) Suppose that we are given $a, b \in \mathbb{Z}$.

Prove that $\forall x, y \in \mathbb{Z}$, $ax + by$ is a multiple of $d = \gcd(a, b)$.

$\gcd(a, b) \mid (ax + by)$

Solution: Let $d = \gcd(a, b)$. Then $d \mid a$ and $d \mid b$. Implies $\exists k, j \in \mathbb{Z}$ such that $a = dk$ and $b = dj$. Thus $ax + by = dkx + djy = d(kx + jy)$ and so $d \mid (ax + by)$ since $kx + jy \in \mathbb{Z}$.

Q2.2 (2 points) 7 divides $4^{186} - 1$. [Hint: $186 = 31 * 6$]

● True ○ False

Solution: True. $4^{186} = (4^6)^{31} \equiv 1^{31} = 1 \pmod 7$ by Fermat's Little Theorem.

Q2.3 (2 points) 5 divides $4^n - 1$ for all **even** $n \in \mathbb{N}$

● True ○ False

Solution: True. $4 \equiv -1 \pmod 5$, so the statement will be true for even n.

Q2.4 (3 points) What is the last digit of the number 3^{47} (i.e. what number 0-9 is in the "one's place")?

7

Solution: The last digit is just the evaluation of $3^{47} \pmod{10}$. Note that $3^2 = 9 \equiv -1 \bmod (10)$, so $3^4 \equiv 1 \pmod{10}$. Then, $3^{47} \equiv 3^{44}3^3 \equiv 1 * 3^3 \equiv 27 \equiv 7 \pmod{10}$.

Q2.5 (3 points) Alice implements RSA correctly with $N = 33$ and $e = 7$ and sends an RSA-encryption $y = 6$ to Bob. Unfortunately, Alice forgot that N should be big because you know that $33 = 3 * 11$. Knowing the encryption y, what was the original message x?

18

Solution:
18.

Knowing $p = 3$ and $q = 11$, we can find $d = e^{-1} (\bmod (p-1)(q-1))$ for $e = 7$. Namely $d = 3 \equiv 7^{-1} (\bmod 20)$. Thus we can compute $y^3 (\bmod 33)$ to recover the original message x. $6^3 = 6^2 * 6 = 36 * 6 \equiv 3 * 6 = 18 (\bmod 33)$.

Q2.6 (4 points) Find $6^7 + 7^6 (\bmod 91)$. **(Hint: use CRT, $91 = 7 * 13$.)**

6

Solution: Answer: As the hint suggests, we can look at the equation in both mod 7 and mod 13 and then combine using CRT.

Let $x = 6^7 + 7^6$.

First we calculate $x \equiv 6^7 + 7^6 \equiv 6^7 + 0 + 6^1 \equiv 6 (\bmod 7)$.

Next, $x \equiv 6^7 + 7^6 \equiv 6^7 + (-6)^6 \equiv 6^7 + 6^6 \equiv 6^6(6+1) = 36^3(6-1) \equiv (-3)^3(7) \equiv -7 \equiv 6 (\bmod 13)$

Thus, since $x \equiv 6$ in both mod 13 and mod 6, then $x \equiv 6 \bmod 13 * 6$ as well by CRT. (Generally, if $x \equiv a (\bmod p)$ and $x \equiv a (\bmod q)$ for coprime p and q, then $x \equiv a (\bmod p * q)$ as well. This was proven in discussion.)

Q3 *Polypourri* **(9 points)**

Q3.1 (3 points) Working in mod 6, how many distinct polynomials of degree *exactly* two are there?

180

Solution: $ax^2 + bx + c$

There are five options for a (in order for the leading coefficient to be nonzero, and thus make sure the polynomial is degree exactly 2), six options for b, and six options for c.

$5 * 6 * 6 = 180$

Q3.2 (3 points) Over GF(7), find an equivalent polynomial $g(x)$ (i.e. outputs the same values when given the same inputs) to $f(x) = 9x^{66} + 7x^{58} + 5x^{55} + 5x^{24} + 10x^{19}$ such that g has degree strictly less than 7 with coefficients from GF(7). That is, it should be the case that $\forall x \in$ GF(7), $f(x) \equiv g(x) \pmod 7$. [Hint: Your solution will be a simple polynomial.]

g(x)=x

Solution: Using Fermat's Little Theorem we know that $x^7 \equiv x \pmod 7$. Thus we can take $f(x)$ and simplify the exponents to $g(x) = 2x^6 + 0 + 5x + 5x^6 + 3x = (2+5)x^6 + (5+3)x \equiv x \pmod 7$ and $f(x) \equiv g(x) \pmod 7$.

Q3.3 (3 points) You receive $n + 2k$ points that come from a degree n-1 polynomial. You know there can be up to k corruptions. How many messages can you recover using Berlekamp–Welch?

1

Solution: Berlekamp–Welch guarantees a unique solution.

Q4 *Graphs Speedrun* **(10 points)**

Q4.1 (2 points) K_4 is planar.

● True ○ False

Solution: True . Doesn't contain K_5 or $K_{3,3}$

Q4.2 (2 points) K_5 has an Eulerian tour.

● True ○ False

Solution: True. Each vertex has degree 4 and so K_5 is connected, even graph.

Q4.3 (2 points) All dice can be viewed as planar graphs wrapped around a sphere. A 12-sided die thus has 12 *faces* to roll. If we know there are 20 pointy *vertices* to 12-sided dice, then how many edges are there on a 12-sided die?

30

Solution:
30. f=12, v=20, and so v+f=e+2 implies that e=30.

Q4.4 (4 points) Prove that if a graph has a vertex that has odd degree, there must be another vertex that has odd degree.

Solution: Contradiction: Assume there is a graph $G = (V, E)$ that has a vertex v' of odd degree but every other vertex has even degree. By the Degree-sum Formula (AKA the Handshake Lemma), $\sum_{v \in V} \deg(v) = 2|E|$, thus the sum of all the vertices' degrees will be a multiple of 2 and thus even. Since adding any amount of even numbers together will still be an even number and deg(v) is even for every $v \neq v'$, we have that $\sum_{v \neq v'} \deg(v)$ is even. Thus, $\sum_{v \in V} \deg(v) = \deg(v') + \sum_{v \neq v'} \deg(v) =$ odd + even = odd. But this contradicts that the handshake lemma tells us that this sum should result in an even number.

Alternate (constructive) proof: We can use the FindTour algorithm from the notes to find another vertex of odd degree: Start at the odd degree vertex and walk to new vertices by crossing any available edge arbitrarily while removing each edge that is crossed until you get stuck.

This algorithm must halt since we remove an edge on each iteration and there are only a finite amount of edges. We will argue that you must get stuck at a different vertex that has odd degree:

We can never get stuck at a vertex that had even degree since entering and leaving such a vertex always deletes edges in pairs (the one you entered on and the one you left on) which always allows you to leave each time you enter due to there being an even number of edges on it.

Further we never get stuck at the starting node because it starts with an odd degree and you start by leaving from there and deleting the edge that was taken, thus leaving an even number of edges attached to it left. Thus we never get stuck there by the same argument made for the vertices that started with an even degree.

Thus we never get stuck at the original odd degree vertex nor at an even degree vertex, yet we must get stuck somewhere since the algorithm must halt and so where it halts must be a new vertex that also has odd degree.

Q5 *Uncountable!?* **(7 points)**

State whether each set is Finite, Countably Infinite, or Uncountably Infinite.

Q5.1 (1 point) The set of all complex numbers $\mathbb{C}$

○ Finite ○ Countably Infinite ● Uncountably Infinite

Solution: Uncountably Infinite, since the set of real numbers (which is uncountably infinite) is a subset of the complex numbers.

Q5.2 (1 point) The set of all prime numbers

○ Finite ● Countably Infinite ○ Uncountably Infinite

Solution: Countably Infinite, since the prime numbers are a subset of the natural numbers (which are countably infinite, thus guaranteeing that the set of prime numbers is countable) and there are infinitely many prime numbers.

Q5.3 (1 point) The set of real numbers between and including 0.001 and 0.1 (i.e the interval [0.001, 0.1])

○ Finite ○ Countably Infinite ● Uncountably Infinite

Solution: Generally, any nontrivial interval on $\mathbb{R}$ will be uncountably infinite. We can establish a bijection $f : [0, 1] \to [0.001, 0.1]$ as $f(x) = 0.001 + 0.099x$. It is invertible with $f^{-1}(y) = \frac{y-0.001}{0.099}$ and thus is a bijection, thus both intervals have the same cardinality. (And we know that the interval $[0, 1]$ is indeed uncountably infinite).

Q5.4 (1 point) The set of complex numbers $a + bi$ such that $a, b \in \mathbb{Z}$

○ Finite ● Countably Infinite ○ Uncountably Infinite

Solution: Notice that this is equivalent to $\mathbb{Z} \times \mathbb{Z}$, which is countably infinite.

Q5.5 (1 point) The set of edges of the complete bipartite graph $K_{100,100}$

● Finite ○ Countably Infinite ○ Uncountably Infinite

Solution: There are exactly $100 * 100$ edges, which is a finite number.

(Question 5 continued...)

Q5.6 (1 point) The set of all integer powers of 2

○ Finite ● Countably Infinite ○ Uncountably Infinite

Solution: We can establish a bijection to the positive integers with the function $f : \{2^n : n \in \mathbb{N}\} \to \mathbb{N}$ as $f(n) = \log_2(n)$ whose inverse is $f^{-1}(k) = 2^k$. The positive integers are countably infinite, thus so is the set of integer powers of 2.

Q5.7 (1 point) The set of all polynomials with degree at most two on $\mathbb{R}$ that pass through $(1, 1)$ and $(2, 2)$

○ Finite ○ Countably Infinite ● Uncountably Infinite

Solution: The polynomials with degree at most 2 are uniquely determined by 3 points. Thus, given 2 points, one more will uniquely determine the polynomial. Given any x value, the y value of the point could be anything in $\mathbb{R}$, and thus is uncountably infinite.

Q6 *All About Compute* **(4 points)**

Two programs F and G are considered *equivalent* if for every input, either both halt with the same output or both don't halt (i.e. $\forall x, F(x) = G(x)$). We will show that deciding whether two arbitrary programs are equivalent is uncomputable:

By contradiction, suppose there exists a procedure *EQUIVALENT(F,G)* that always halts and returns TRUE iff programs F and G are equivalent. We want to construct a program *HALT(P,x)* that decides whether program P halts on input x, thus solving the Halting Problem.

Fill in the blanks below to implement *HALT* using *EQUIVALENT*.

```
def HALT(P,x):

    def F(y):
        ----(1)----
        ----(2)----

    def G(y):
        return 0

    return EQUIVALENT(----(3)----, ----(4)----)
```

1:$P(x)$

2:return 0

3:F

4:G

Solution: (1): P(x) (2): return 0 (3): F or G (4): G or F, respectively

This makes it so that if $P(x)$ halts, then F will always halt and return 0, thus *EQUIVALENT*(F,G) will return True. Elsewise, if $P(x)$ fails to halt (infinite loops), then *EQUIVALENT*(F,G) will return False. This indeed solves the Halting Problem, which is impossible, and thus *EQUIVALENT* cannot exist.

Q7 *Let's Count!* **(18 points)**

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 Σ or Π)

Q7.1 (3 points) A card deck has 52 cards; each of four suits has thirteen ranks (A, 2–10, J, Q, K). A "poker hand" is an unordered set of five cards. How many poker hands contain exactly one 6 and one 7?

$4 * 4 * \binom{44}{3}$

Solution: Choose which 6 is in the hand (4 choices) and which 7 is in the hand (4 choices). The remaining 3 cards can be anything except a 6 or 7, so we choose them from the remaining 44 cards.

Q7.2 (3 points) Suppose we draw 3 cards without replacement. Let A be the event that we draw at least one ace. Let C_1, C_2, and C_3 be the event that we draw exactly one ace, exactly two aces, and exactly three aces, respectively. Let F, S, and T be the events that the first, second, and third card drawn was an ace. Is it easier to write $P(A)$ in terms of C_i or F, S, T? Explain why your choice is easier.

- ● C_i
- ○ F, S, T

Solution: We can see that $P[A] = P[C_1] + P[C_2] + P[C_3]$ since the C_i are mutually exclusive and their union is A. However, $P[A] = P[A_1 \cup A_2 \cup A_3] = P[A_1] + P[A_2] + P[A_3] - P[A_1 \bigcap A_2] - P[A_1 \bigcap A_3] - P[A_2 \bigcap A_3] + P[A_1 \bigcap A_2 \bigcap A_3]$.

Agnes is distributing 40 indistinguishable girl scout cookies to her 3 minions, Bob, Kevin, and Stuart, to help her sell them! However, each minion can only carry at most 15 cookies.

Q7.3 (3 points) If there were no carrying limit, how many ways are there of distributing the 40 cookies?

$\binom{40+2}{2}$

Solution: This is counting problem 4. Recall that this says that, to split n objects into k groups, there are $\binom{n+k-1}{k-1}$ ways to do so. In this case, $n = 40$ and $k = 3$.

Q7.4 (3 points) How many ways are there of distributing the 40 cookies such that Bob's limit is exceeded?

$\binom{26}{2}$

Solution: Give 16 cookies to Bob first, then apply counting problem 4 on the remaining 26 cookies among 3 minions.

Q7.5 (2 points) How many ways are there such that both Bob's and Kevin's limits are exceeded?

$\binom{10}{2}$

Solution: Give 16 cookies to both Bob and Kevin first, then apply counting problem 4 on the remaining 8 cookies.

Q7.6 (4 points) How many ways are there to distribute the 40 cookies while respecting the limits of the minions? Give your answer terms of a, b, and c, which are the correct answers to 7.3, 7.4, and 7.5.

$a - 3b + 3c$

Solution: Use inclusion–exclusion. Start with all distributions, subtract those where at least one minion exceeds their limit, then add back those where two limits are exceeded. (All three cannot be exceeded since $n < 3k$.)

$\binom{42}{2} - 3 * \binom{26}{2} + 3 * \binom{10}{2}$

Q8 *The Paradoxical Survey* **(7 points)**

In a given town, $\frac{3}{4}$ of the residents have black hair and $\frac{2}{3}$ are female. Having black hair and being female are independent. Suppose the mayor selects two people at random from the town with replacement.

The mayor tells Manuel that at least one of the two people is a black haired female.

Manuel is told to guess whether both people have black hair. Given the information from the mayor, what is the probability that both people have black hair? Make sure to show your work.

Hint: The answer is not $\frac{9}{16}$.

Answer: $\frac{2}{3}$

Solution: Let A be the event that at least one of the two people is a black haired female. Let B be the event that both people have black hair.

First find the probability that a randomly chosen person has black hair and is female:

$$P(\text{black haired female}) = \left(\frac{3}{4}\right) * \left(\frac{2}{3}\right) = \frac{1}{2}.$$

Then

$$P(A) = 1 - P(\text{neither people are black haired female}) = 1 - \left(1 - \frac{1}{2}\right)^2 = \frac{3}{4}.$$

The probability that both people have black hair is

$$P(B) = \left(\frac{3}{4}\right)^2 = \frac{9}{16}.$$

Now compute $P(A \mid B)$. Given that both people have black hair, this becomes the probability that at least one of the two people is female. Each person independently has a $\frac{2}{3}$ chance of being female, so

$$P(A \mid B) = 1 - P(\text{neither is female} \mid B) = 1 - \left(1 - \frac{2}{3}\right)^2 = 1 - \left(\frac{1}{3}\right)^2 = \frac{8}{9}.$$

By Bayes' rule,

$$P(B \mid A) = \frac{P(A \mid B) * P(B)}{P(A)} = \frac{\left(\frac{8}{9}\right) * \left(\frac{9}{16}\right)}{\frac{3}{4}} = \frac{2}{3}.$$

Q9 *Flipping Random Variables* **(13 points)**

Flip a biased coin infinitely many times. Each flip is heads with probability p, tails otherwise, independently of others. For each of the following situations, give the name of the distribution and its parameters. The first is provided for you as an example. Recall the common discrete distributions in this class are Bernoulli, Binomial, Geometric, Uniform, and Poisson.

Q9.1 (0 points) First three flips are heads.

$$X \sim \text{Bernoulli}(p^3)$$

Q9.2 (2 points) Number of heads in the first 70 flips.

$$X \sim \text{Bin}(70, p)$$

Solution: This is the classical example of the binomial distribution, where we count the number of successful trials out of n, where each independently has a probability p of success.

Q9.3 (2 points) If there's exactly one head in the first 20 flips, the flip number where the first head occurs.

$$X \sim \text{Unif}\{1, 2, ..., 20\}$$

Solution: Given that there's exactly one head in the first 20 flips, the only possible outcomes are HTT...T, THTT...T, TTHT...T, i.e. all the outcomes with exactly 1 head and 19 tails. Each outcome has the same probability $p^1(1-p)^{19}$ of occurring, so they are all equally as likely. Thus, each position from 1 to 20 is equally likely to be the position with the single heads. This exactly describes the discrete uniform distribution on the set $\{1, 2, ..., 20\}$.

In other words, symmetry!

Q9.4 (3 points) Gap between first two tails. For example, the sequence THHHT... would have a gap of length 3. This distribution is some standard random variable minus 1.

$X \sim \text{Geo}(1-p) - 1$

Solution: After getting the first tail, we are essentially counting the number of heads until we get the next tail. Since the Geometric distribution is memoryless, this is equivalent to counting the number of heads until we get the next tail from an arbitrary starting point. The total number of flips until we get a tail is $\text{Geo}(1-p)$, and since the final flip is a tail, we subtract 1 to only count the number of heads.

Q9.5 (3 points) Number of flips until the same flip occurs twice in a row. Assume for this problem that $p = \frac{1}{2}$. This distribution is 1 plus some standard random variable.

$X \sim 1 + \text{Geo}\left(\frac{1}{2}\right)$

Solution: Notice that after the first flip, every subsequent flip has a 0.5 chance of being the same as the previous flip. An example can help illustrate this: imagine we flip a heads first. Then, it is a 0.5 chance to subsequently flip a heads and and the experiment. Elsewise, if we flip a tails, and then the next flip has a 0.5 chance to be tails, etc. Thus, after the first flip, each subsequent flip has a 0.5 chance of success, which exactly describes a $\text{Geo}\left(\frac{1}{2}\right)$ RV.

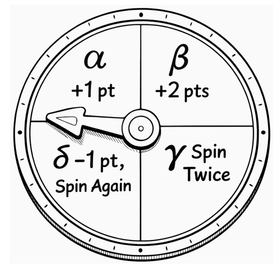
Q9.6 (3 points) Number of runs in the first 50 flips. Define a "run" as an uninterrupted sequence of the same flip. For example if the flips are `HHHHHTTTHTTH...`, we have the run `HHHHH`, followed by the run `TTT`, followed by the run `H`, followed by the run `TT`, ... Assume for this problem that $p = \frac{1}{2}$. This distribution is 1 plus some standard random variable.

$X \sim 1 + \text{Bin}\left(49, \frac{1}{2}\right)$

Solution: Similarly to the previous question, we start off with a flip to begin the sequence. For every subsequent flip (for which there are 49 of), it has a 0.5 chance of breaking the current run, and thus starting a new run. The latter 49 flips can then be described as $\text{Bin}\left(49, \frac{1}{2}\right)$, and we add 1 for the starting flip.

Q10 *The Wheel* **(3 points)**

JQ loves the kids menu at the CS70 Restaurant, which has a game on the back. Each player gets exactly one spin on a wheel, though some of the options on the wheel allow additional spins. The player's score is the sum of the results of their spin. The possible outcomes for each spin are equally likely and are listed below:



α: Earn 1 point.
β: Earn 2 points.
γ: Spin two additional times.
δ: Lose 1 point and spin again.

Examples:

1. The player gets α on their first spin, so they earn 1 point and the game is over.
2. The player gets δ on their first spin, so their score becomes −1. The player gets γ, so now they get to spin twice. On the first of these spins, they get γ again, so increase their remaining spins to three. The player gets α, then β, then α, for a total score of −1 + 1 + 2 + 1 = 3. Overall spin sequence: $\delta\gamma\gamma\alpha\beta\alpha$.

Let X be the score of a player. What is $E[X]$?

○ 0 ○ 1 ● 2 ○ 3 ○ 4 ○ ∞

Solution: Let X be the expected points of a single spin. We can set up a recurrence relation, similar to the Markov Chain hitting time equations.

$E[X] = \frac{1}{4}(1) + \frac{1}{4}(2) + \frac{1}{4}(2E[X]) + \frac{1}{4}(E[X] - 1)$

Solving yields $E[X] = 2$.

Q11 *Summing* **(7 points)**

Q11.1 (3 points) Suppose $X \sim \text{Normal}(1, 2)$ and $Y \sim \text{Normal}(2, 1)$ are independent random variables. What is the distribution of $Z = 2X - Y + 1$? State its name and specify its parameter(s).

$Z \sim \text{Normal}(1, 9)$

Solution: Sum of independent normal distributions will still be normal (and adding a constant to a normal distribution just shifts the mean), so we calculate $E[Z] = 2 * 1 - 2 * 1 + 1 = 1$ and $\text{Var}(Z) = 2^2 * 2 + (-1)^2 * 1 + 0 = 9$, thus $Z \sim \text{Normal}(1, 9)$.

Q11.2 (2 points) Suppose we have two independent random variables $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\lambda)$. Let $Z = X + Y$. Which of the following are true?

- [x] $Z \sim \text{Poisson}(2\lambda)$
- [ ] $Z = 2P$, where $P \sim \text{Poisson}(\lambda)$
- ○ None of the above

Solution: A property of Poisson distributions is that if $X \sim \text{Poisson}(\lambda_x)$ and $Y \sim \text{Poisson}(\lambda_y)$ are independent, then $X + Y \sim \text{Poisson}(\lambda_x + \lambda_y)$. The first option directly follows from this. The second option is incorrect because $Z = 2P$ would not be distributed as $\text{Poisson}(2\lambda)$; in fact, it will no longer be a Poisson RV. The easiest way to see this is that $E[Z] = 2E[P] = 2\lambda$ but $\text{Var}(Z) = 2^2\text{Var}(P) = 4\lambda$

Q11.3 (2 points) Suppose we have n independent random variables $X_i \sim \text{Poisson}(\lambda)$. Let $S = \sum_{i=1}^{n} X_i$. Which of the following are true? Check all that apply.

- [x] For all n, $S \sim \text{Poisson}(\lambda n)$
- [x] In the large n limit, $\frac{S-\lambda n}{\sqrt{\lambda n}}$ converges to $\text{Normal}(0, 1)$
- ○ None of the above

Solution: The first option follows from the independence property of Poissons described in the previous question. The second option follows from CLT, since $\frac{S-\lambda n}{\sqrt{\lambda n}}$ has mean of 0 and variance of 1 and S is a sum of i.i.d. RVs.

Q12 *The Waiting Game* **(12 points)**

Sam and Andy are waiting for students in their office hours. They've observed that students arrive at a rate of one student every 30 minutes, and decide to model students arrivals as a Poisson random variable.

Q12.1 One day, nobody has shown up after 50 minutes. They're debating whether they should leave early to eat pineapple pizza. What is the probability that at least one student arrives in the last ten minutes?

$1 - e^{-\frac{1}{3}}$

Solution: Let $\lambda_{10} = \frac{1}{3}$. Let S_1 be the number of students who arrive in ten minutes. Then $S_1 \sim$ Poisson(λ_{10}).

$P[S_1 = 0] = e^{-\lambda_{10}} = e^{-\frac{1}{3}}$, so $P[S_1 > 0] = 1 - e^{-\frac{1}{3}}$.

Q12.2 Suppose Jay observes the same rate of student arrivals. However, due to his excellent teaching skills, once a student arrives at his OH, they do not leave.

If nobody is in the room at the start of the hour, what is the expected number of students in the room at the end of the hour?

2

Solution: Let $\lambda_{60} = 2$. Let S_2 be the number of students who arrive in one hour. Then $S_2 \sim$ Poisson(λ_{60}) and $E[S_2] = \lambda_{60} = 2$.

Suppose we model the departure of BART trains as an exponential random variable. On average, BART trains leave Richmond every 15 minutes.

Q12.3 If Ishan shows up at a random time, what is his expected wait for a BART train to leave?

15 minutes

Solution: 15 minutes. By the memoryless property, this is equal to the expected wait time for a departure from any arbitrary point.

Q12.4 A station agent informs Ishan that the most recent train left 5 minutes ago; what is his expected wait for a BART train?

15 minutes

Solution: 15 minutes. Same as the previous part, since whatever happens in a previous time interval is irrelevant for a memoryless RV.

Q12.5 For this part, trains arrive at some rate λ_t trains per hour. Passengers arrive at the station at some rate λ_p people per hour. Once passengers arrive, they take the first train that leaves.

Let A_p represent the first arrival time for a passenger and A_t represent the first arrival time for a train. The joint PDF of these exponential random variables is:

$$f_{A_p,A_t}(a_p, a_t) = \lambda_p \lambda_t e^{-\lambda_p a_p - \lambda_t a_t}.$$

for $a_p, a_t \geq 0$ and 0 elsewise.

Write an expression for $P(A_t < A_p)$ but do not solve; this is the probability that a train arrives before any passengers.

$$P(A_t < A_p) = \int_0^\infty \lambda_t e^{-(\lambda_p+\lambda_t)a_t} \, \mathrm{d}a_t = \frac{\lambda_t}{\lambda_p + \lambda_t}.$$

Solution:

$$P(A_t < A_p) = \int_0^\infty \left(\int_{a_t}^\infty \lambda_p \lambda_t e^{-\lambda_p a_p - \lambda_t a_t} \, \mathrm{d}a_p \right) \mathrm{d}a_t.$$

Alternatively, we can opt to use total probability rule, noting that:

$f_{A_t}(a_t) P(a_t < A_p) = \lambda_t e^{-\lambda_t a_t} \int_{a_t}^\infty e^{-\lambda_p a_p} \, \mathrm{d}a_p = \lambda_t e^{-\lambda_t a_t} e^{-\lambda_p a_t} = \lambda_t e^{-(\lambda_t+\lambda_p)a_t}.$

and thus

$$P(A_t < A_p) = \int_0^\infty \lambda_t e^{-(\lambda_p+\lambda_t)a_t} \, \mathrm{d}a_t = \frac{\lambda_t}{\lambda_p + \lambda_t}.$$

Q13 *Bingo* **(17 points)**

At the MLK Student Union (bound) Open House, every booth in the building is represented by exactly one square on a 5×5 bingo card (25 squares total). For each booth, Melody decides **independently** whether to visit it. She visits any given booth with probability p and skips it with probability $1 - p$. A square is **marked** if Melody visits its booth.

There are 12 possible **bingo lines**: the 5 rows, the 5 columns, and the 2 diagonals. Let X denote the total number of bingo lines that are completely marked. Note that there is no "free square".

Q13.1 (3 points) Compute the expectation $E[X]$.

$12 * p^5$

Solution: Let X_i be the indicator r.v. for the event that line i (a row, column, or diagonal) is completely marked. Then

$$X = \sum_{i=1}^{12} X_i$$

and by linearity of expectation,

$$E[X] = \sum_{i=1}^{12} E[X_i] = 12 * p^5,$$

since a line has 5 squares and all must be visited.

Q13.2 (2 points) Consider two different bingo lines that share exactly one square. What is the probability that both get completely marked?

p^9

Solution: If they share a square, then a total of 9 squares need to be filled between the two of them, each of which has an independent p probability of being filled.

Q13.3 (2 points) Consider two different bingo lines that do not share any squares. What is the probability that both get completely marked?

p^{10}

Solution: If they don't share a square, then a total of 10 squares need to be filled between the two of them, each of which has an independent p probability of being filled.

Q13.4 (5 points) Let a be the answer to Q13.1, let b be the answer to Q13.2, and let c be the answer to Q13.3.

Compute the variance $\mathrm{Var}(X)$. Feel free to leave your answer in terms of a,b, and/or c.

$$12p^5 + 2 * (20p^{10} + 46p^9) - (12p^5)^2$$

Solution:

Consider pairs of bingo lines. If they are both rows or both columns, they share no squares: $2 * \binom{5}{2} = 20$ such pairs. Any other pair shares exactly one square, so the number of sharing pairs is

$$\binom{12}{2} - 20 = 46.$$

For a fixed line i,

$$E[X_i] = p^5.$$

For $i \neq j$,

$$E[X_i X_j] = p^{10} \quad \text{if lines i, j share no squares,}$$

$$E[X_i X_j] = p^9 \quad \text{if lines i, j share one square.}$$

Now expand:

$$E[X^2] = E\left[\left(\sum_{i=1}^{12} X_i\right)^2\right] = \sum_{i=1}^{12} E[X_i] + 2 * \sum_{1 \le i < j \le 12} E[X_i X_j].$$

Grouping by overlap type gives

$$E[X^2] = 12p^5 + 2 * (20p^{10} + 46p^9).$$

Also $E[X] = 12p^5$, so

$$\mathrm{var}(X) = E[X^2] - (E[X])^2 = 12p^5 + 2 * (20p^{10} + 46p^9) - (12p^5)^2.$$

In terms of the previous parts, this would be $a + 92b + 40c - a^2$. (Equivalent simplified form: $\mathrm{var}(X) = 12p^5 + 92p^9 - 104p^{10} - 144p^{10} + 40p^{10}$ is messy; keeping the grouped expression above is usually cleanest.)

Q13.5 (2 points) Students win a sticker if they have at least 1 bingo line. Give an upper bound for the probability that Melody achieves at least one bingo line, i.e. $P[X \geq 1]$, using the union bound.

$12 * p^5$

Solution: Let L_i be the event that line i is completely marked. Then

$$P[X \geq 1] = P(\cup_{i=1}^{12} L_i) \leq \sum_{i=1}^{12} P(L_i) = 12p^5.$$

Q13.6 (3 points) Students win a pair of socks if they have at least 2 bingo lines. Give an upper bound for the probability that Melody achieves at least two bingo lines, i.e. $P[X \geq 2]$, using Chebyshev's inequality. Give your answer in terms of $E[X]$ and $\text{Var}(X)$.

$\frac{\text{Var}(X)}{(2-E[X])^2}$

Solution: Apply Chebyshev to the deviation above the mean:

$$P[X \geq 2] = P[X - E[X] \geq 2 - E[X]] \leq \frac{\text{Var}(X)}{(2 - E[X])^2},$$

assuming $2 > E[X]$ (otherwise the bound is trivial).

Q14 *Covariance Proof* **(5 points)**

Consider two events A, B with nonzero probability and indicators $\boldsymbol{I}_A, \boldsymbol{I}_B$ for them.

Show that, if $\operatorname{Cov}(\boldsymbol{I}_A, \boldsymbol{I}_B) > 0$, then $P(A \mid B) > P(A)$.

Solution:

$$\operatorname{Cov}(\boldsymbol{I}_A, \boldsymbol{I}_B) = E[\boldsymbol{I}_A * \boldsymbol{I}_B] - E[\boldsymbol{I}_A] * E[\boldsymbol{I}_B] = P(A \wedge B) - P(A)P(B)$$

Also, $P(A \bigcap B) = P(A \mid B)P(B)$, so

$$\operatorname{Cov}(\boldsymbol{I}_A, \boldsymbol{I}_B) = P(A \mid B)P(B) - P(A)P(B) = P(B) * (P(A \mid B) - P(A)).$$

Since $P(B) > 0$ and the covariance is positive, we must have $P(A \mid B) - P(A) > 0$, hence $P(A \mid B) > P(A)$.

Q15 *Continuous Random Variable Potpourri* **(13 points)**

Q15.1 (2 points) Suppose we have continuous random variables $X \sim \text{Gaussian}(3, 6)$, $Y \sim \text{Exponential}(3)$. What is $P(X = Y)$?

$P(X = Y) =$ 0

Solution: Generally, two continuous random variables will be equal to each other with 0 probability, unless they have some predetermined dependency like $X = Y$ or $X = B * Y + (1 - B) * Z$ where B is a Bernoulli random variable.

Concretely, we can use law of total probability: $P[X = Y] = \int_0^\infty P[X = Y|Y = y] f_Y(y) dy = \int_0^\infty 0 f_Y(y) dy = 0$

Q15.2 (2 points) For $X \sim \text{Gaussian}(3, 6)$, which of the following is true about $P(X < 4)$?

- ○ < 0.5
- ○ $= 0.5$
- ● > 0.5
- ○ Not enough information

Solution: For a Gaussian with mean μ, exactly half of the probability distribution is below μ, and half is above μ. That is, $P[X < \mu] = P[X > \mu] = 0.5$.

Then, $P[X < 4] = P[X < 3] + P[3 \le X < 4] > P[X < 3] = 0.5$.

Suppose we have three independent continuous random variables over the interval between 0 and 12, i.e. $A, B, C \sim$ Unif$(0, 12)$. Let X be an indicator that C is between A and B.

Q15.3 (2 points) What is $P(X = 1|\ A, B)$? In other words, if we know A and B, what is the probability that C is between them? Give your answer in terms of A and B and potentially the absolute value function.

$P(X = 1|A, B) = \frac{|A-B|}{12}$

Solution: Given the values of A, B, the probability that C is between them is the probability that C lands in the interval $[A, B]$ (or $[B, A]$ in the case that B is bigger. Either way, the interval has length $|A - B|$). Given that any value between $[0, 12]$ is equally as likely, the probability is thus $\frac{|A-B|}{12}$

Q15.4 (2 points) Suppose we don't know A and B. What is $P(X = 1)$? In other words, if we pick three random numbers in the range $[0, 12]$, what is the chance that the third one we pick is between the other two?

Hint: There's a way to do this without any integrals.

$P(X = 1) = \frac{1}{3}$

Solution: Since all orderings of A, B, C are equally likely, by symmetry

$$P(X = 1) = \frac{1}{3}.$$

Q15.5 (5 points) What is $E[|A - B|]$? Hint: There's a way to solve this problem using the previous two parts that avoid the need for any integrals. Warning: This problem is particularly challenging.

$E[|A - B|] = 4$

Solution:

We know from earlier that if we don't know A and B, then X is true with probability 1/3. However, if we know A and B, then X is true with probability $\frac{|A-B|}{12}$.

That is, imagine running an experiment where we pick an A, B, and C in that order and check to see if C lands in the middle. Before we pick A and B, we know there's a 1/3 chance that C will land in the middle. But after we pick A and B, this chance will depend on what we got for A and B, e.g. if we picked 1.5 and 11.5, C will land between them with probability $\frac{10}{12}$.

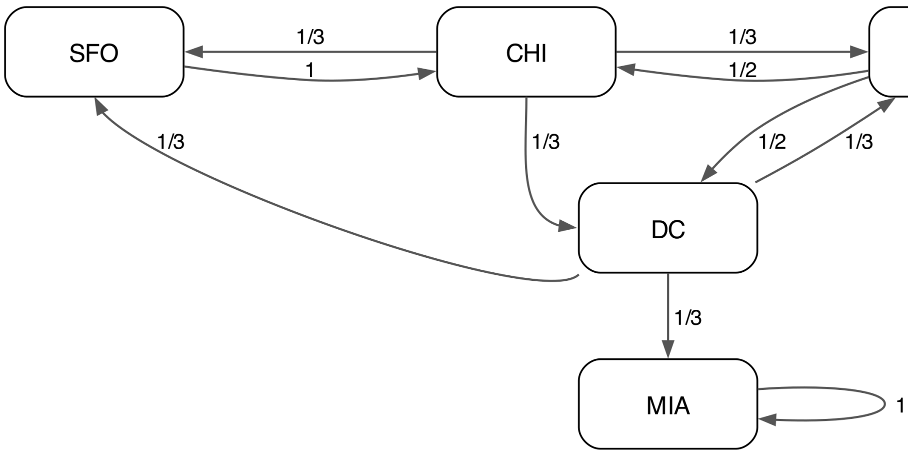
Naturally, if we were to run these experiments over and over, we'd see that on average C would land in the middle 1/3 of the time.

That is, on average $\frac{|A-B|}{12}$ is equal to 1/3, or equivalently $E\left[\frac{|A-B|}{12}\right] = \frac{1}{3}$. Multiplying both sides by 12 we have $E[|A - B|] = 4$.

More formally, we can also write this in terms of our iterated expectation formula, namely $E[E[X \mid A, B]] = E[X]$. Here, we already know that $E[X] = \frac{1}{3}$ and $E[E[X \mid A, B]] = E[P(X = 1 \mid A, B)] = E\left[\frac{|A-B|}{12}\right]$.

Q16 *Markov Trains* **(6 points)**

Tom is exploring America using trains. He's attempting to make it from Chicago (CHI) to Miami (MIA). However, America's trains are incredibly unreliable. Whenever Tom boards a train, he does not know its destination. The probability of a train being bound for a station is uniformly distributed among the outgoing routes.



Q16.1 (3 points) Tom is trying to avoid running into his uncle in New York City (NYC). What is the probability that Tom will make it to Miami from Chicago without visiting NYC?

Write out a system of equations that you could use to find the answer, but do not solve the equations.

Solution: Let $\alpha(x)$ be the probability of reaching "MIA" before ever visiting "NYC", starting from x.

Boundary conditions:
- $\alpha(\text{MIA}) = 1$
- $\alpha(\text{NYC}) = 0$

Recurrences:

$$\alpha(\text{SFO}) = \alpha(\text{CHI})$$

$$\alpha(\text{DC}) = \frac{1}{3}\alpha(\text{MIA}) + \frac{1}{3}\alpha(\text{NYC}) + \frac{1}{3}\alpha(\text{SFO}) = \frac{1}{3} + \frac{1}{3}\alpha(\text{CHI})$$

$$\alpha(\text{CHI}) = \frac{1}{3}\alpha(\text{SFO}) + \frac{1}{3}\alpha(\text{DC}) + \frac{1}{3}\alpha(\text{NYC}) = \frac{1}{3}\alpha(\text{CHI}) + \frac{1}{3}\left(\frac{1}{3} + \frac{1}{3}\alpha(\text{CHI})\right)$$

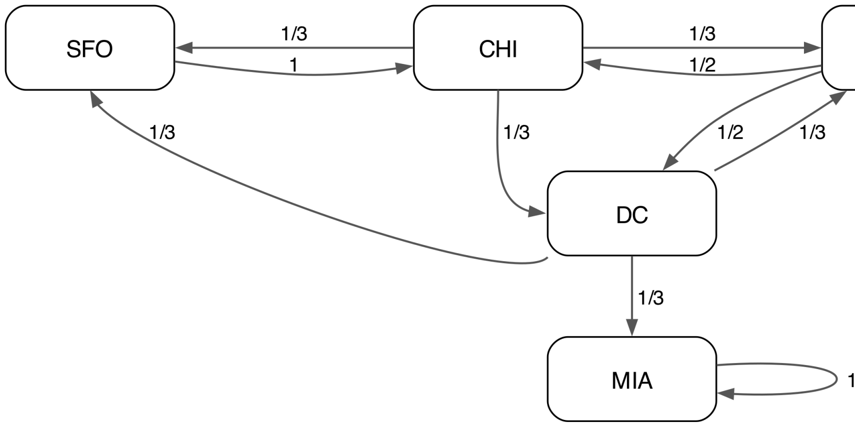
So

$$\alpha(\text{CHI}) = \frac{4}{9}\alpha(\text{CHI}) + \frac{1}{9}$$

So

$$\alpha(\text{CHI}) = \frac{4}{9}\alpha(\text{CHI}) + \frac{1}{9} \Rightarrow \frac{5}{9}\alpha(\text{CHI}) = \frac{1}{9} \Rightarrow \alpha(\text{CHI}) = \frac{1}{5}.$$

Q16.2 (3 points)



Tom needs this trip to be over before his finals week; however his uncle doesn't intend to make that easy. Every time Tom ends up in NYC, his uncle will force him to stay there for three days. In all other cities, Tom will spend one day there before moving on to his next destination.

Assume train travel is instant and takes zero days.

What is the expected number of days Tom takes to arrive in Miami (starting from CHI and including the initial day spent there)? Write out a system of equations that you could use to find the answer, but do not solve the equations.

Solution: Let $\beta(x)$ be the expected number of **days until first arrival in "MIA"**, starting from x. Set $\beta(\text{MIA}) = 0$.

Recurrences (add the "time spent in the city after arriving"):

$$\beta(\text{CHI}) = 1 + \frac{1}{3}\beta(\text{SFO}) + \frac{1}{3}\beta(\text{DC}) + \frac{1}{3}\beta(\text{NYC})$$

$$\beta(\text{SFO}) = 1 + \beta(\text{CHI})$$

$$\beta(\text{DC}) = 1 + \frac{1}{3}\beta(\text{MIA}) + \frac{1}{3}\beta(\text{NYC}) = 1 + \frac{1}{3}\beta(\text{NYC})$$

$$\beta(\text{NYC}) = 3 + \frac{1}{2}\beta(\text{CHI}) + \frac{1}{2}\beta(\text{DC})$$