CS 70 Discrete Mathematics and Probability Theory
Summer 2025 Tate

Final Solutions

PRINT Your Name: Oski Bear

SIGN Your Name: $\mathscr{O}\mathscr{S}\mathscr{K}\mathscr{I}$

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) What is the probability that **at least one** student answers "1" to this question?

Answer: 1: If you say 1, then at least one student has answered 1!

3. **Propositional Logic**

1. $(\neg P \implies P) \equiv P$

   Answer: True: If you convert the implication to the equivalent "OR" form, you get $P \vee P$, which is equivalent to P.

2. $((P \vee Q) \wedge (P \implies R) \wedge (Q \implies R)) \equiv R$

   Answer: False: If P and Q are both False and R is True, the left hand side evaluates to False (because the $P \vee Q$ clause is False), but the right side is True.

3. $\exists x(P(x) \wedge \neg Q(x)) \equiv \forall x(\neg P(x) \implies Q(x))$

   Answer: False: It's pretty easy to find counter-examples to this. For example $P(x) =$"x is odd" and $Q(x) =$"x is prime" (the left hand size is true, and the right hand size is false).

4. $\exists x(P(x) \wedge \neg Q(x)) \equiv \neg \forall x(P(x) \implies Q(x))$

   Answer: True: Convert the implication in the right hand side into an OR, and then push the negation all the way in (past the quantifiers and into the OR) using De Morgan's Law, and you'll get the left hand side.

5. What is the simplest (least number of symbols) proposition equivalent to $(P \wedge Q) \vee (P \wedge \neg Q)$?

   Answer: P

6. What is the simplest (least number of symbols) proposition equivalent to $P \wedge (P \implies Q)$?

   Answer: $P \wedge Q$ – note that it's tempting to say "Q" here, but that's not right! Think about the case when P is False and Q is True, making the expression False while Q is True.

4. Basic Proof Techniques

1. The following predicate is part of the basis for a specific proof technique – what is it?
$\bigwedge_{n=1}^{k} P(n) \implies P(k+1)$
Answer: Strong Induction. Partial credit for just "induction."

2. (6 points) Let $x, y \in \mathbb{R}$. Prove that if $\frac{x+y}{2} = a$, then either $x \leq a$ or $y \leq a$.
Answer: We prove this by contraposition. If $x > a$ and $y > a$, then $x + y > 2a$ and $\frac{x+y}{2} > a$. That means $\frac{x+y}{2} \neq a$, completing the proof of the contrapositive.

5. Graphs

1. Any graph in which every vertex has degree at most 2 is planar.
Answer: True: Such a graph must consist of connected components each of which is either a simple path or a single simple cycle. Such components (and graphs) are obviously planar.

2. Any graph that has a vertex of degree 6 is not planar.
Answer: False: If *all* vertices have degree 6 then the graph can't be planar, but think of a tree where the root has degree 6 (and all trees are planar).

3. A simple cycle with n vertices is bipartite iff n is odd.
Answer: False: It's bipartite iff n is even!

4. Let $\mathbb{H}^d$ denote the d-dimensional hypercube.
   (a) What is the minimum number of colors needed for a valid vertex-coloring of $\mathbb{H}^d$, the d-dimensional hypercube?
   Answer: 2: all hypercubes are bipartite (in particular, if you consider the binary strings labeling vertices, you can put all with an even number of 1's on one side and an odd number of 1's on the other – since every edge flips one bit in the label, it only connects even numbers of ones to odd numbers of ones and vice versa). $d = 0$ is a degenerate case that only requires 1 color, but in general 2 colors are needed.
   (b) (8 points) Let G be a connected planar graph with at least 2 edges and no cycles of length 3. If n is the number of vertices, what is the maximum amount of edges? *Show all of your work that justifies your bound!*
   [Hint: Use Euler's formula, along with a bound on the number of faces.]
   Answer: Let's call every instance where a face touches an edge a "side" (of the face), and with $e \geq 2$ and no cycles of length 3, every face in a planar embedding of G must have at least 4 sides. Using the e, v, f variables from Euler's formula, this means the number of sides is $\geq 4f$. We can also count sides by edges, and since every edge has two sides there are $2e$ sides. Combining these two ways of counting sides, we have $f \leq e/2$. Euler's formula says $e = v + f - 2$, so $e \leq v + e/2 - 2$ which implies that $e \leq 2v - 4$. Putting this back into the notation from the problem, with n vertices, the maximum number of edges is $\mathbf{2n - 4}$.
   (c) (4 points) Prove that the 4-dimensional hypercube is non-planar.
   Answer: $\mathbb{H}^4$ has 2^4 vertices and $(4 \cdot 2^4)/2 = 32$ edges. All hypercubes are bipartite (see part (a)), and bipartite graphs cannot have odd-length cycles (so no cycles of length 3). That means that for a hypercube to be planar it must satisfy the bound in part (b), but in this case $2v - 4 = 28$, and there are more than 28 edges. Therefore, $\mathbb{H}^4$ is not planar. (Note that $\mathbb{H}^3$ is though!)

6. Stable Matching

1. In a stable matching instance with $n \geq 2$, if every candidate lists the same job last in their preference list, then the propose-and-reject algorithm (with jobs proposing) cannot terminate in a single day.

   Answer: False: If every job has a different favorite candidate, then the algorithm terminates in one day regardless of candidate orderings.

2. If there is objective criteria for ranking candidates, so all jobs use the same ranking, then there is only one stable matching.

   Answer: True. This can be seen inductively. The base case (one job and one candidate) is trivial. For the induction step, let C_1 be the candidate at the top of all job lists, and let J_1 be that candidate's top-ranked job. If J_1 is not matched to C_1 in some matching, then they are a rogue pair (since J_1 prefers C_1 and C_1 prefers J_1). Therefore C_1 and J_1 are matched in any stable matching (this is just a specific instance of the general principle given in question 4 below). Now remove J_1 and C_1 from the problem, and we're left with an instance of size $n-1$ where every job has the same candidate ranking, so by the induction hypothesis there is only one stable matching for those remaining $n-1$ jobs/candidates.

3. In a job optimal stable matching, it is impossible for every candidate to get their favorite job.

   Answer: False: If job J_i has top candidate is C_i (for all i), and candidate C_i's favorite job is J_i, then the propose and reject algorithm (which finds the job-optimal stable matching) will match up every pair (J_i, C_i) and every candidate is matched to their favorite job.

4. If Job J lists candidate C as their top choice, and candidate C lists job J as their favorite job, then J and C are matched in any stable matching.

   Answer: True. If J and C are not matched, then they are a rogue pair and the matching is not stable.

7. Modular Arithmetic and Friends

1. The following two questions refer to this system of equations:

$$\begin{cases} x \equiv 1 \pmod{10} \\ x \equiv 2 \pmod{15} \end{cases}$$

   (a) There is a value of x that satisfies these constraints.
   Answer: False.

   (b) (4 points) Give a one sentence justification of your answer.
   Answer: Looking at the equations mod 5, they imply that $x \equiv 1 \pmod 5$ and $x \equiv 2 \pmod 5$, which is impossible.

2. The following two questions refer to this system of equations:

$$\begin{cases} x \equiv 6 \pmod{10} \\ x \equiv 1 \pmod{15} \end{cases}$$

   (a) There is a value of x that satisfies these constraints.
   Answer: True.

   (b) (4 points) Give a one sentence justification of your answer.
   Answer: $x = 16$ satisfies the system of equations.

3. Prove the following two claims.
   (a) (6 points) Prove that for any modulus m, $x^2 \equiv (m-x)^2 \pmod{m}$.
   Answer: Consider the difference of the two sides: $(m-x)^2 - x^2 = (m^2 - 2mx + x^2) - x^2 = m^2 - 2mx = m(m-2x)$. Since the difference is a multiple of m, the sides are equivalent mod m.
   (b) (8 points) Define $S_n = \sum_{k=1}^{n} (-1)^k \cdot k^2$. Prove that if $n \geq 2$ is even, then S_n is a multiple of $(n+1)$.
   Answer: (Note that there's a solution using closed-form for the sum too, and possibly other proof methods.)
   We let $m = n+1$ and look at S_n mod m. Pairing up terms from the front and end, we consider the kth term (for $k \leq n/2$) and the $n+1-k$th term. The $n+1-k$th (or $m-k$th) term is $(-1)^{n+1-k}(m-k)^2 \equiv (-1)^{n+1-k}k^2 \pmod{m}$ (by part (a)). Furthermore, since n is even, k is even if and only if $n+1-k$ is odd, so these paired terms have equal value and opposite signs and cancel out in the summation. Since there are an even number of terms, each one is paired with a value that cancels it out, so $S_n \equiv 0 \pmod{m}$. Therefore, S_n is a multiple of $m = n+1$.

8. Polynomials, Secret Sharing, and Error-Correcting Codes

1. For this problem, all operations are in $GF(5)$. Given $x, y \in GF(5)$, how many polynomials of degree 3 go through (x,y)?
   Answer: Accepted answers were 100 and 125. Unfortunately, this problem was stated vaguely. In the context of secret sharing, when we say something like "select a random polynomial of degree 3" then it typically means "degree 3 or less" as the leading coefficient of the random polynomial might be zero. There are 125 such polynomials. If the "degree 3" is taken exactly, then there are only 100 polynomials since the cubic term cannot have a zero coefficient. The exam question should have been more explicit, rather than leaving the "typical usage" vs "strict mathematical definition" vague.
2. A secret s has been shared using the secret sharing scheme over GF(5) with a degree 2 polynomial $P(x)$. Bryan receives share $(1,3)$, Pratyay: $(2,4)$, Altay: $(4,2)$.
   (a) What is the secret?
   Answer: 4. The secret sharing polynomial is $P(x) = ax^2 + bx + c$, and we can set up the following system of equations to find the coefficients:
   $a + b + c \equiv 3 \pmod{5}$
   $4a + 2b + c \equiv 4 \pmod{5}$
   $a + 4b + c \equiv 2 \pmod{5}$

   Solving, we find that $P(x) = x^2 + 3x + 4$, so the secret is $P(0) = 4$.
   (b) If Kyle received a share of the form $(3,y)$, what is y?
   Answer: 2. Evaluating $P(3)$ over the integers, we get 22. Taking this mod 5 we get the answer.
3. A message is sent using Reed-Solomon coding with polynomial $P(x)$ over $GF(11)$, transmitting packets $P(1), P(2), \ldots, P(8)$
   (a) If the code can correct for up to 2 corrupted packets, what is the length of the message (in packets)?
   Answer: 4. Using Reed-Solomon for a message of length n and allowing for corrections of k packets, we send $n+2k$ packets. We know from the problem that $k = 2$ and 8 packets were sent, so $n = 4$.
   (b) If a message is received and the Berlekamp-Welch algorithm computes polynomial $E(x) = x^2 - 6x + 8$, which packets are corrupted?
   Answer: Positions 2 and 4 (or you could say $P(2)$ and $P(4)$ were corrupted). Factoring $E(x)$ we get $E(x) = (x-2)(x-4)$, and the error positions are at the roots of this polynomial.

9. Counting

For the questions in this section, you may leave your answer as an expression using factorial notation, $n!$, or the choose notation, $\binom{n}{k}$, unless otherwise specified.

In this question, Minqing is playing CS70Nite, the latest, greatest computer game.

1. The first level of the game in played in rounds, each of which results in a win or loss. If he wins the round, he earns 30 points, and if he loses the round, he loses 30 points (point totals can be negative).

   Minqing plays a total of 10 rounds. How many distinct sequences of wins and losses are possible such that he finishes the 10 rounds with a total score of +60 points?

   Answer: $\binom{10}{4}$ (or $\binom{10}{6}$). The only way to play 10 rounds and end with +60 points is to win 6 games and lose 4. There are $\binom{10}{4}$ ways to choose which 4 games are lost.

2. Once completing the first level, Minqing wants to set up an account. He needs to select a 10-digit ID, where all digits can be values 0-9 but no two consecutive digits can be the same. How many different IDs could he pick?

   Answer: $10 \cdot 9^9$. There are 10 choices for the first digit, and there are 9 choices for each of the subsequent 9 digits.

3. After setting up an account, Minqing needs to distribute 8 points among three stats: Cardinality, Discreteness, and Computability. Cardinality is non-negative, and the others are positive. How many ways can he distribute the points among the three stats?

   Answer: $\binom{8}{2}$. This can be converted into a basic "stars and bars" problem: Let x_1 be the points for cardinality, and x_2 and x_3 be the points for Discreteness and Computability *minus one*, respectively. Now we have three non-negative variables which should add up to 6, so this is a stars and bars problem with 6 stars (for the sum of the variables) and 2 bars (to separate the three variables). There are 8 stars/bars positions, and so there are $\binom{8}{2}$ ways to place the bars.

4. Now Minqing must round out his trio by selecting a captain and a lieutenant from the the remaining 11 TAs. How many ways can he do this?

   Answer: $11 \cdot 10$ (or 110). Note that the roles are distinct, so it's *not* just selecting $\binom{11}{2}$.

5. There are 7 game locations that he can visit, each with a different item. He must collect at least two items to finish the game. How many possible collections of items can he finish the game with?

   Answer: $2^7 - 8$ (or 120). There are 2^7 possible subsets of items, and the one empty subset or the 7 subsets with a single item are insufficient to finish the game.

6. In the next part of the game, there are two islands he can visit, with n checkpoints on each of them. To go between locations, he can only take the BijectionBus, which can to to any location on the other island (it will not go between locations on the same island, however). He can take the BijectionBus as many times as he likes. How many different paths could he take to visit each checkpoint exactly once and end at the same location he started?

   Answer: $n!(n-1)!$. Starting out, there are n locations on the other island to choose from. Then you come back to one of $n-1$ unvisited locations on "island 1," and in your third move have $n-1$ locations to choose from back on "island 2." Continuing this and multiplying the number of choices together we have $n \cdot (n-1) \cdot (n-1) \cdot (n-2) \cdot (n-2) \cdot (n-3) \cdot (n-3) \cdots$. The choices corresponding to the "other island" are $n!$, and on the starting island are $(n-1)!$.

7. Minqing has lost track of time and needs to start his homework! Help him by answering this homework question: How many functions from $\{1,2,\ldots,n\}$ to $\{1,2,\ldots,n\}$ are not bijections?

   Answer: $n^n - n!$. There are n^n possible functions, and $n!$ of them are bijections.

10. Countability and Computability

1. The set of functions from $\mathbb{N}$ to $\{\text{True}, \text{False}\}$ is countable.

   Answer: False

2. The set of finite length binary strings is countable.

   Answer: True

3. A "halting analyzer" is a program that takes another program as input, and determines whether the input program halts when it is run. It is impossible to write a halting analyzer that works correctly for some input programs.

   Answer: False. We can certainly decide whether *some* programs halt (a program with no loops or recursion, for example). We just can't write an analyzer that decides correctly for *all* programs.

4. A "love for CS70 analyzer" is a program that takes another program as input, and determines whether the input program ever prints "I love CS70!" when it is run. It is impossible to write a love of CS70 analyzer that works for all input programs.

   Answer: True. This is the same (just a different string) as your discussion problem showing that no analyzer can decide whether a program prints "Hello world."

5. A "bounded execution error detector" takes a program and a number N as input, and determines whether the input program encounters a runtime error in the first N steps of its execution. It is impossible to write a bounded execution error detector that works for all inputs.

   Answer: False. This analyzer is in fact easy to write. Just run the program for N steps and see if it encounters a runtime error!

6. (8 points) Variables in Python can be assigned values with types that are determined at runtime, and a variable can even hold different values of different types at different times during the execution of a program. For example, the variable `x` could be assigned a string value and then later assigned an integer value.

   You have been asked to write an analyzer program whose input is a Python program and the name of a variable, and returns "True" if the program ever assigns a string to the named variable when run and "False" otherwise. Prove that it's impossible to write such an analyzer that works correctly on all inputs.

   Answer: Assume for the sake of contradiction that such an analyzer exists, and `A(P,"variablename")` always correctly returns `True` if P ever assigns a string to `variablename`, and `False` if it doesn't. We use this to create a program `Halt(P,x)` that solves the halting problem, returning `True` if program `P` halts when run on `x`, and `False` otherwise:

```
def Halt(P,x):
   Scan P for variable names, and let newvarname be an unused name
   Create program P2 as follows:
       def P2():
             newvarname = 0
             P(x)
             newvarname = "a string"
   return A(P2,"newvarname")
```

   `P2` assigns a string to `newvarname` if and only if `P(x)` halts, and since `A` correctly detects this condition, our function `Halt` correctly solves the halting problem. No such program is possible, giving us a contradiction, so no analyzer program can detect if a program saves a string to a particular variable.

11. Discrete Probability Basics

Jia Jun rolls an *unfair* 5-sided die, with sides numbered 1 through 5. The die has a probability $p = 1/2$ of rolling a 1, and all other numbers are equally likely. Let A be the event that the number rolled is odd, and let B be the event that the number rolled is a perfect square.

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

   Answer: 3/4. For this problem, we can make a table with each roll, its probability, and whether the roll (outcome) is in each event:

Roll	Prob	A	B
1	1/2	✓	✓
2	1/8		
3	1/8	✓	
4	1/8		✓
5	1/8	✓	

   Finding $\mathbb{P}[A]$ is then just a matter of adding up probabilities in rows with A checked.

2. What is $\mathbb{P}[\overline{B}]$?

   Answer: 3/8. Referring to the table in (a), add up rows in which B is not checked.

3. What is $\mathbb{P}[A \cup B]$?

   Answer: 7/8. Referring to the table in (a), add up rows in which either A or B is checked.

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

   Answer: 4/5. The conditional probability formula says that $\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$, and adding up probabilities from the table for numerator and denominator we get $\frac{1/2}{5/8} = 4/5$.

5. What is $\mathbb{P}[A \cap B \mid A \cup B]$?

   Answer: 4/7. Using the conditional probability formula like in the previous question (and note that $(A \cap B) \cap (A \cup B) = A \cap B$), we get $\frac{1/2}{7/8} = 4/7$.

6. What is $\mathbb{P}[A \cap \overline{B} \mid B]$?

   Answer: 0. You don't need to evaluate any formula here, just note that any outcome in A and $\overline{B}$ can't be in B.

7. (5 points) If you can re-design the die to change the probability p, subject to $p < 1$, what value of p will make A and B are independent?

   Answer: 1/3. Using a generic value for p, we find the following probabilities: $\mathbb{P}[A] = p + (1-p)/2$, $\mathbb{P}[B] = p + (1-p)/4$, and $\mathbb{P}[A \cap B] = p$. To be independent, we need $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$, which leads to the quadratic $3p^2 - 4p + 1 = 0$ or $(3p-1)(p-1) = 0$. This has solutions at $p = 1/3$ and $p = 1$, and since $p < 1$ is a condition this leaves just $p = 1/3$.

12. Bayes for Plagiarists

Note: For this problem, gives answers as fractions in lowest terms. Do not divide out to give decimal representation.

The Stanfurd School of Cosmetology, Programming, and Stuff has launched a new plagiarism detector to flag suspicious assignment submissions. If a submission is plagiarized, it gets flagged with probability p_1. If a submission is not plagiarized, it gets flagged with probability p_2.

Suppose that in CS 1, 10% of submissions are plagiarized, whereas in CS 100, 2% of submissions are plagiarized.

1. Is it true that $p_1 + p_2 = 1$?

   Answer: Depends. The sum can be 1 (as in the part below), but there's no reason it must be.

2. Let $p_1 = 0.8, p_2 = 0.2$. Suppose that a submission is flagged. Given this information, what is the probability that it was actually plagiarized...

   (a) ...for a CS 1 submission?

   Answer: 4/13. With the sample space being the set of CS1 submissions, we let F be the event that a submission was flagged, and P be the event that a submission was plagiarized. We know that $\mathbb{P}[F|P] = 0.8$, $\mathbb{P}[F|\overline{P}] = 0.2$, and $\mathbb{P}[P] = 0.1$, and the problem is asking for $\mathbb{P}[P|F]$. We can calculate $\mathbb{P}[F] = \mathbb{P}[F|P] \cdot \mathbb{P}[P] + \mathbb{P}[F|\overline{P}] \cdot \mathbb{P}[\overline{P}] = 0.8 \cdot 0.1 + 0.2 \cdot 0.9 = .26$. We now compute $\mathbb{P}[P|F] = \frac{\mathbb{P}[F|P] \cdot \mathbb{P}[P]}{\mathbb{P}[F]} = \frac{0.8 \cdot 0.1}{0.26} = \frac{8}{26} = \frac{4}{13}$.

   (b) ...for a CS 100 submission? *Hint: they're not the same!*

   Answer: 4/53. This is the same calculation as in the last part, but with $\mathbb{P}[P] = 0.02$. From this we calculate $\mathbb{P}[F] = \mathbb{P}[F|P] \cdot \mathbb{P}[P] + \mathbb{P}[F|\overline{P}] \cdot \mathbb{P}[\overline{P}] = 0.8 \cdot 0.02 + 0.2 \cdot 0.98 = 0.016 + 0.196 = 0.212$. We then compute $\mathbb{P}[P|F] = \frac{\mathbb{P}[F|P] \cdot \mathbb{P}[P]}{\mathbb{P}[F]} = \frac{0.8 \cdot 0.02}{0.212} = \frac{16}{212} = \frac{4}{53}$.

3. An auditor comes in and does the following experiment: First one of the two classes is randomly selected (probability $1/2$ of each), and then a random (uniformly distributed) submission is selected from that class. The plagiarism detector is not used at all in this experiment. After careful examination, the auditor determines that the submission was plagiarized. Given the probabilities above, what is the probability that CS 1 was the selected class?

   Answer: 5/6. Let X be a random variable indicating the class that was sampled, so $\mathbb{P}[X = 1] = 1/2$ and $\mathbb{P}[X = 100] = 1/2$. We want to know what $\mathbb{P}[X = 1 \,|\, P]$ is. We first compute $\mathbb{P}[P] = \mathbb{P}[P \,|\, X = 1] \cdot \mathbb{P}[X = 1] + \mathbb{P}[P \,|\, X = 100] \cdot \mathbb{P}[X = 100] = 0.1 \cdot 0.5 + 0.02 \cdot 0.5 = .06$. Then $\mathbb{P}[X = 1 \,|\, P] = \frac{\mathbb{P}[P \,|\, X=1] \cdot \mathbb{P}[X=1]}{\mathbb{P}[P]} = \frac{0.1 \cdot 0.5}{.06} = \frac{5}{6}$.

13. Senators of the Round Table

It's the year 2462 and the United States still exists, but there are now $n \geq 2$ states. Each state still has two senators, with a "junior senator" and a "senior senator" from each state. There is a reception being held in which all $2n$ senators will attend, and they are randomly assigned seats around a large round table (every seat being equally likely for each senator).

1. California still exists! What is the probability that the two senators from California are seated together?
   Answer: $\frac{2}{2n-1}$. We'll calculate this with a simple counting argument (there are easier arguments, but this works for the more complicated problems below, so it's a good warm-up): There are $(2n)!$ possible seating arrangements. If the two senators are sitting together, we can always decide which one is to the left (since $n \geq 2$). The leftmost California senator can be either the junior or senior senator (2 choices) and is in one of $2n$ seats. The remaining $2n-2$ senators can be arranged in $(2n-2)!$ ways, so the total number of ways the two senators from California can be seated together is $2 \cdot 2n \cdot (2n-2)!$. Therefore, the probability of this happening is $\frac{2n \cdot 2 \cdot (2n-2)!}{(2n)!} = \frac{2}{2n-1}$.
2. (6 points) New Mexico still exists! What is the probability that the two senators from California are seated together given that the two senators from New Mexico are seated together?
   Answer: $\frac{2}{2n-2}$. If C is the event that the two senators from California are seated together, and N is the event that the two senators from New Mexico are seated together, we will compute $\mathbb{P}[C|N] = \frac{|C \cap N|}{|N|}$. The number of ways the two senators from New Mexico can be seated together is the same calculation we did in question 1 above for California, so $|N| = 2 \cdot 2n \cdot (2n-2)!$.

   If senators from both California and New Mexico are seated together, the leftmost senator from California can be in one of $2n$ positions, but then the leftmost senator from New Mexico can be in only $2n-3$ positions (they can't be immediately to the left or right of the leftmost California senator). There are 4 junior/senior orderings of the senators from these states, and there are $(2n-4)!$ orderings for the remaining senators. Therefore there are $|C \cap N| = 4 \cdot 2n \cdot (2n-3) \cdot (2n-4)!$.

   Combining our results, we get $\mathbb{P}[C|N] = \frac{4 \cdot 2n \cdot (2n-3) \cdot (2n-4)!}{2 \cdot 2n \cdot (2n-2)!} = \frac{2}{2n-2}$.
3. (6 points) The senior senator from California (David Yang the 18th) has a degree in Mechanical Engineering, as does a senator from New Mexico (these are the only two mechanical engineers in the senate). Senator Yang is happy if he is seated next to the other senator from California or the other mechanical engineer, and he is unhappy otherwise. What is the probability that Senator Yang is happy?
   Answer: $\frac{8n-10}{(2n-1)(2n-2)}$ – it's OK if this is given as $\frac{4}{2n-1} - \frac{2}{(2n-1)(2n-2)}$ without simplifying.
   The number of ways for the senior senator to be seated next to the junior senator (event J) is $2 \cdot 2n \cdot (2n-2)!$, as we've calculated above. This is the same as the number of ways for the senior senator to be seated next to the other mechanical engineer (event M). We want $|J \cup M|$, but these events are not disjoint so we must use the principle of inclusion-exclusion and subtract the size of the intersection $J \cap M$. In $J \cap M$ either the junior senator is to the left of the senior senator and the mechanical engineer is to the right, or vice-versa. That's two possibilities, and then there are $(2n-3)!$ ways to arrange the remaining senators, so $|J \cap M| = 2 \cdot 2n \cdot (2n-3)!$.
   Combining these results we get $|J \cup M| = 4 \cdot 2n \cdot (2n-2)! - 2 \cdot 2n \cdot (2n-3)! = (8n-10) \cdot 2n \cdot (2n-3)!$. Dividing by the total number of seating arrangements, $(2n)!$, we get the final answer.
4. Let N_i be the event that senator i is seated next to the other senator from their state.
   (a) For any senator, what is $\mathbb{P}(N_i)$?
   Answer: $\frac{2}{2n-1}$
   (b) Are the N_i events pairwise independent?
   Answer: No.

(c) Give a brief justification of your previous answer.

Answer: The simplest argument is that if the two senators from California are senators 0 and 1, then $\mathbb{P}[N_0 \mid N_1] = 1 \neq \mathbb{P}[N_0]$ (in other words, knowing that the junior senator is seated next to the senior senator implies that the senior senator is seated next to the junior senator!). You could also point out that the probabilities in questions 1 and 2 are different.

(d) What is the expected number of senators seated next to the other senator from their state?

Answer: $\frac{4n}{2n-1}$. Let I_i be an indicator variable for event N_i, and from part (a) above we see that $\mathbb{E}[I_i] = \frac{2}{2n-1}$. The number of senators seated next to the other senator from their state is $I_1 + I_2 + \cdots + I_{2n}$, and while these indicator variables are not independent we use linearity of expectation to give the expected value of this sum: $2n\frac{2}{2n-1} = \frac{4n}{2n-1}$.

14. Declaration of Independence

Consider the following theorem:

Theorem: Given n pairwise independent random variables $X_1, X_2, \ldots, X_n$, it's true that

$$\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n).$$

1. Ijin's brother NotIjin wants to prove this theorem, and as a first step has proven the following lemma:

   Lemma: If X_1 and X_2 are independent, then $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2)$.

   NotIjin loves induction so decides to prove the general theorem with induction, and includes this in his induction step:

   Given $X_1, X_2, \ldots, X_n$, we define $Y = X_1 + X_2 + \cdots + X_{n-1}$, and apply the induction hypothesis to get $\mathrm{Var}(Y) = \mathrm{Var}(X_1 + X_2 + \cdots + X_{n-1}) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_{n-1})$. Now we can apply the lemma to get

   $$\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(Y + X_n) = \mathrm{Var}(Y) + \mathrm{Var}(X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n).$$

   Unfortunately, this is not logically correct. Identify the specific error and what is wrong with it.

   Answer: The X_i's are pairwise independent, not mutually independent, so Y and X_n might not be independent.

2. (8 points) Ijin is much smarter than his brother, and can write a correct proof. Be like Ijin, and write a correct proof of the theorem.

   Answer: We expand and use linearity of expectation to calculate

   $$\mathbb{E}[(X_1 + X_2 + \cdots + X_n)^2] = \mathbb{E}[\sum_{i=1}^{n} X_i^2 + \sum_{i \neq j} X_i X_j] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i X_j] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i]\mathbb{E}[X_j],$$

   where the last part ($\mathbb{E}[X_i X_j] = \mathbb{E}[X_i]\mathbb{E}[X_j]$) only requires pairwise independence. Next we calculate

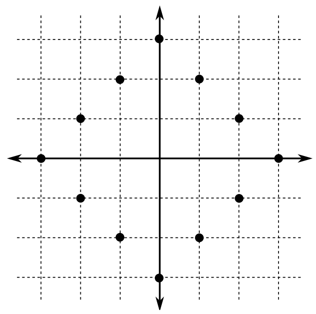
   $$(\mathbb{E}[(X_1 + X_2 + \cdots + X_n)])^2 = (\mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]))^2 = \sum_{i=1}^{n} \mathbb{E}[X_i]^2 + \sum_{i \neq j} \mathbb{E}[X_i]\mathbb{E}[X_j],$$

   Finally,

   $$\begin{aligned}
   \mathrm{Var}(X_1 + X_2 + \cdots + X_n) &= \mathbb{E}[(X_1 + X_2 + \cdots + X_n)^2] - (\mathbb{E}[(X_1 + X_2 + \cdots + X_n)])^2 \\
   &= \sum_{i=1}^{n} \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i]\mathbb{E}[X_j] - \sum_{i=1}^{n} \mathbb{E}[X_i]^2 - \sum_{i \neq j} \mathbb{E}[X_i]\mathbb{E}[X_j] \\
   &= \sum_{i=1}^{n} \mathbb{E}[X_i^2] - \sum_{i=1}^{n} \mathbb{E}[X_i]^2 \\
   &= \sum_{i=1}^{n} \left(\mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2\right) \\
   &= \sum_{i=1}^{n} \mathrm{Var}(X_i)
   \end{aligned}$$

15. Point Probabilities

1. A square is drawn in the plane with corners at $(0,3)$, $(3,0)$, $(0,-3)$, and $(-3,0)$, and we let S be the set of points on the perimeter with integer coordinates. See the picture below, where the 12 points in S are marked:



We uniformly choose a point from some subset of S.

(a) (4 points) We uniformly choose from the entire set S, and let X be a random variable giving the x coordinate and Y be a random variable giving the y coordinate. What are the values below?

Answer: $\mathbb{E}[X] = 0$, $\mathbb{E}[Y] = 0$, $\text{Var}[Y] = 19/6$, and $\text{cov}(X,Y) = 0$.

(b) Are X and Y independent?
Answer: No.

(c) Give a brief (but mathematically precise) justification for your previous answer.
Answer: If $X = 0$, then Y must be 3 or -3. Specifically, $\mathbb{P}[X = 0, Y = 2] = 0$, but $\mathbb{P}[X = 0] = 1/6$ and $\mathbb{P}[Y = 2] = 1/6$ so $\mathbb{P}[X = 0] \cdot \mathbb{P}[Y = 2] = 1/36 \neq 0$.

(d) (4 points) Define S_1 to be just the 7 points with non-negative y coordinates. We choose a point uniformly from S_1 and let X_1 be a random variable giving the x coordinate and Y_1 be a random variable giving the y coordinate. What are the values below?

Answer: $\mathbb{E}[X] = 0$, $\mathbb{E}[Y] = 9/7$, $\text{Var}[Y] = 52/49$, and $\text{cov}(X,Y) = 0$.

(e) (4 points) Define S_2 to be just the 4 points with non-negative x and y coordinates. We choose a point uniformly from S_2 and let X_2 be a random variable giving the x coordinate and Y_2 be a random variable giving the y coordinate. What are the values below?

Answer: $\mathbb{E}[X] = 3/2$, $\mathbb{E}[Y] = 3/2$, $\text{Var}[Y] = 5/4$, and $\text{cov}(X,Y) = -5/4$.

16. Distributions and Bounds

1. Let $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ be independent.
   (a) Find $\mathbb{P}[X+Y=k]$.
   Answer: $\binom{n+m}{k}p^k(1-p)^{n+m-k}$
   (b) What is distribution of $X+Y$?
   Answer: $X+Y \sim \text{Bin}(n+m, p)$
   (c) (5 points) Give an intuitive interpretation for why $X+Y$ has the distribution you identified in part (b), without using formulas for probabilities.
   Answer: $\text{Bin}(n, p)$ is the distribution of the number of successes in n independent Bernoulli trials with success probability p. Y just adds m more independent Bernoulli trials with the same success probability, so we have a total of $n+m$ independent Bernoulli trials with success probability p. That's $\text{Bin}(n+m, p)$.
2. Let X denote the number of heads you obtain after flipping a fair coin 100 times.
   (a) Use Chebyshev's inequality to bound the probability that either $X \leq 10$ or $X \geq 90$.
   Answer: $X \sim \text{Bin}(100, 1/2)$, so $E[X] = 50$ and $\text{Var}(X) = 25$. Chebyshev's inequality says that $\mathbb{P}[|X-\mu| \geq c] \leq \frac{\text{Var}(X)}{c^2}$, and plugging in $c = 40$ we get $\mathbb{P}[|X-50| \geq 40] \leq \frac{25}{40^2} = \frac{1}{64}$
   (b) For this distribution, $\mathbb{P}[X \leq 10] = \mathbb{P}[X \geq 90]$.
   Answer: True. The distribution is symmetric around $\mu = 50$.
   (c) Use the central limit theorem to estimate the probability from part (a). Your answer can include $\mathbb{P}[Z \geq c]$ for some c, where $Z \sim N(0, 1)$ (in other words, Z is a standard normal variable).
   Answer: $2\mathbb{P}[Z \geq 8]$
   X is the sum of $n = 100$ Bernoulli random variables with mean $\mu = 1/2$ and standard deviation $\sigma = 1/2$, and the Central Limit Theorem says that $\frac{X-n\mu}{\sqrt{n}\sigma} \sim N(0, 1)$, so $\mathbb{P}[\frac{X-n\mu}{\sqrt{n}\sigma} \geq c]$ is approximately $\mathbb{P}[Z \geq c]$.
   Here $\sqrt{n}\sigma = 5$ and $n\mu = 50$, so $\mathbb{P}[\frac{X-50}{5} \geq c] = \mathbb{P}[X \geq 50+5c]$ is approximately $\mathbb{P}[Z \geq c]$, and with $c = 8$ we get $\mathbb{P}[X \geq 90]$ is approximately $\mathbb{P}[Z \geq 8]$. Then $\mathbb{P}[X \leq 10 \text{ or } X \geq 90]$ is approximately $2\mathbb{P}[Z \geq 8]$.
   Note that while there's no way to have calculated this on the exam, this probability is approximately 1.244×10^{-15}.

17. Second Chances in Gambling

Consider the following dice game: You initially pay a casino \$8 to roll 2 fair six-sided dice. After you roll once, you may re-roll one or both dice in a second round of rolling, subject to varying rules described below. After the second round (if used) the casino pays you \$$s$, where s is the sum of the values showing on the dice. Your "winnings" are s minus your cost to play. For example, if you paid the original \$8 to play (and no more), and after the second round the dice show a 4 and a 6, then $s = 10$ and your winnings are \$2. Note that winnings can be negative.

1. You are not given the opportunity to re-roll in the second round. What are your expected winnings?
   Answer: -1. Why: The two dice are independent, so we just focus on one and let R be a random variable representing the final value on one die. $E[R] = 7/2$, so the expected winnings are $2E[R] - 8 = 2(7/2) - 8 = -1$.
2. In game variant 1, you will re-roll any die that came up "1" in your original roll. You don't need to pay anything extra to do this. What is your expected winnings in this game?

Answer: $-1/6$. Why: Looking at one die, let R_1 be a random variable indicating the value of the first roll, and R be the final value on the die. Then $E[R] = E[R \mid R_1 > 1] \cdot \mathbb{P}[R_1 > 1] + E[R \mid R_1 = 1] \cdot \mathbb{P}[R_1 = 1]$. If $R_1 > 1$ then the value is uniformly distributed over $2, \ldots 6$, with expected value 4 (this happens with probability $5/6$). If $R_1 = 1$ then we re-roll and the expected value of the re-roll is $7/2$. Therefore, $E[R] = 4 \cdot (5/6) + (7/2) \cdot (1/6) = 47/12$ and the expected winnings are $2(47/12) - 8 = -1/6$.

3. (4 points) In game variant 2, the casino allows you to select a value c in advance, and then you re-roll any die roll that was $\leq c$ in the first roll. What value of c will maximize your expected winnings, and what is your expected winnings with that c?

Answer: $c = 3$; expected winnings 1/2.
$E[R] = E[R \mid R_1 > c] \cdot \mathbb{P}[R_1 > c] + E[R \mid R_1 \leq c] \cdot \mathbb{P}[R_1 \leq c]$. Using reasoning like in the previous part, we have $E[R \mid R_1 > c] = \frac{7+c}{2}$, and with the other values we have

$$\begin{aligned} E[R] &= E[R \mid R_1 > c] \cdot \mathbb{P}[R_1 > c] + E[R \mid R_1 \leq c] \cdot \mathbb{P}[R_1 \leq c] \\ &= \frac{7+c}{2} \cdot \frac{6-c}{6} + \frac{7}{2} \cdot \frac{c}{6} \\ &= \frac{(7+c)(6-c)+7c}{12} \\ &= \frac{42+6c-c^2}{12} \end{aligned}$$

To find the c that maximizes the expected roll, we take the derivative of the numerator and set to 0 to find $c = 3$ (we can and should also verify that this gives a local maximum). So the best strategy is to re-roll any time the first roll is 3 or smaller. The expected roll is then $\frac{42+18-9}{12} = \frac{51}{12}$ and the expected winnings are $2 \cdot \frac{51}{12} - 8 = \frac{3}{6} = \frac{1}{2}$.

4. (4 points) Game variant 3 is the same as game variant 2, except that the casino will charge you \$1 for each re-roll (you can only re-roll each die once, however). What value of c will maximize your expected winnings, and what is your expected winnings with that c?

Answer: $c = 2$; expected winnings -1/3.
The effect of charging a dollar for a re-roll effectively decreases the payout by \$1 if we re-roll, which can capture by reducing $E[R \mid R_1 \leq c]$ by 1. Doing the calculations again, we get $E[R] = \frac{42+4c-c^2}{12}$. Taking the derivative and solving for c we get $c = 2$, and at this value we have $E[R] = \frac{46}{12} = \frac{23}{6}$. Now our winnings are $2 \cdot \frac{23}{6} - 8 = -\frac{1}{3}$.

5. The casino decides they really don't like people re-rolling in the second round. They decide to charge \$d for a re-roll so that any gambler that re-rolls any die will only decrease their expected winnings. What is the smallest integer value of \$d that the casino can use?

Answer: $d = 4$.
Adjusting the previous calculations to take into account charging \$d for a re-roll, we get $E[R] = \frac{42+(6-2d)c-c^2}{12}$ and the casino wants to set d so that this is $E[R] < 7/2$ (the expected initial roll). To do this, the casino needs

$$\frac{42+(6-2d)c-c^2}{12} < \frac{7}{2} \iff 42+(6-2d)c-c^2 < 42 \iff (6-2d)c-c^2 < 0$$

We guarantee this by selecting $d > 3$, and since the problem says d has to be an integer the casino should charge \$4 for a re-roll.

18. Network Traffic

1. Two systems are sending packets to a receiver, with the packet from system one arriving at time $T_1 \sim \text{Exp}(\lambda_1)$ and the packet from system two arriving at time $T_2 \sim \text{Exp}(\lambda_2)$. Let R be a random variable representing the time at which the first packet (from either system) arrives.

   (a) What is $f_R(x)$, the PDF for R?
   Answer: $(\lambda_1 + \lambda_2)e^{-(\lambda_1+\lambda_2)x}$. The first arrival $R = \min(T_1, T_2)$, and this is exponentially distributed with parameter $\lambda_1 + \lambda_2$.

   (b) What is the expected value for R?
   Answer: $\frac{1}{\lambda_1+\lambda_2}$

   (c) Write out an definite integral that is used to compute $\mathbb{P}[T_1 \leq T_2]$.
   Answer:

   $$\int_0^\infty \int_0^y \lambda_2 e^{-\lambda_2 y} \lambda_1 e^{-\lambda_1 x} \, \mathrm{d}x \, \mathrm{d}y$$

   (d) Evaluate the integral from the previous part to give $\mathbb{P}[T_1 \leq T_2]$.
   Answer: $\frac{\lambda_1}{\lambda_1+\lambda_2}$

2. Two systems are sending packets, as above but with arrival times $T_1 \sim \text{Uniform}([0,10])$ and $T_2 \sim \text{Uniform}([0,10])$. Let S be a random variable representing the time at which the first packet arrives.

   (a) What is $f_S(x)$, the PDF for S? You only need to give the function when $0 \leq x \leq 10$.
   Answer: $f_S(x) = \frac{10-x}{50}$.
   $\mathbb{P}[S > x] = \mathbb{P}[T_1 > x] \cdot \mathbb{P}[T_2 > x] = \frac{(10-x)^2}{100}$
   So $\mathbb{P}[S \leq x] = 1 - \frac{(10-x)^2}{100}$, and taking the derivative to get the PDF we have $f_S(x) = \frac{10-x}{50}$.

   (b) What is the expected value for S?
   Answer: $\frac{10}{3}$

   $$\int_0^{10} x \cdot \frac{10-x}{50} \, \mathrm{d}x = \int_0^{10} \left(\frac{1}{5}x - \frac{1}{50}x^2\right) \mathrm{d}x = \left(\frac{1}{10}x^2 - \frac{1}{150}x^3\right)\Bigg|_0^{10} = \frac{100}{10} - \frac{1000}{150} = \frac{10}{3}.$$

   (c) (5 points) What is the probability that the two packets arrive within one time unit of each other (in other words, $\mathbb{P}[|T_1 - T_2| \leq 1]$)?
   Answer: $\frac{19}{100}$.
   This is basically the "Lunch Meeting" problem from the discussion. We calculate the area of the width 2 strip around the diagonal, which is 19, so $\mathbb{P}[|T_1 - T_2| \leq 1] = \frac{19}{100}$.

19. Apply the Clamps

A *clamping function* $g_c(x)$ ensures that a value does not exceed some maximum allowable value:

$$g_c(x) = \begin{cases} x & \text{if } x \le c; \\ c & \text{otherwise.} \end{cases}$$

1. (8 points) Let $X \sim \text{Exp}(\lambda)$. Let Y be a clamped version, with $Y = g_{2/\lambda}(X)$ (so Y will never be greater than twice the mean of X). Derive $\mathbb{E}[Y]$, showing your work (leave constants such as e in symbolic form – do not try to give a decimal representation):

   Answer: We set up the standard integral for expected value, and then separate into two parts at $x = 2/\lambda$ (where $g_c(x)$ changes):

$$\mathbb{E}[Y] = \int_0^\infty g_{2/\lambda}(x) \cdot \lambda e^{-\lambda x}\,\mathrm{d}x = \int_0^{2/\lambda} x\lambda e^{-\lambda x}\,\mathrm{d}x + \int_{2/\lambda}^\infty \frac{2}{\lambda} \cdot \lambda e^{-\lambda x}\,\mathrm{d}x$$

   The first part can be solved using integration by parts, where $u = x$ and $v' = \lambda e^{-\lambda x}\,\mathrm{d}x$ (so $u' = \mathrm{d}x$ and $v = -e^{\lambda x}$):

$$\begin{aligned}
\int_0^{2/\lambda} x\lambda e^{-\lambda x}\,\mathrm{d}x &= x \cdot (-e^{-\lambda x})\Big|_0^{2/\lambda} + \int_0^{2/\lambda} e^{-\lambda x}\,\mathrm{d}x \\
&= x \cdot (-e^{-\lambda x})\Big|_0^{2/\lambda} + \left(-\frac{1}{\lambda}e^{-\lambda x}\right)\Big|_0^{2/\lambda} \\
&= -\frac{2}{\lambda}e^{-2} - 0 - \frac{1}{\lambda}e^{-2} + \frac{1}{\lambda} \\
&= \frac{1}{\lambda} - \frac{3}{\lambda}e^{-2}
\end{aligned}$$

   The integral from $2/\lambda$ to ∞ does not require integration by parts:

$$\int_{2/\lambda}^\infty \frac{2}{\lambda} \cdot \lambda e^{-\lambda x}\,\mathrm{d}x = -\frac{2}{\lambda}e^{-\lambda x}\Big|_{2/\lambda}^\infty = \frac{2}{\lambda}e^{-2}$$

   Combining these results we get the expected value:

$$\mathbb{E}[Y] = \frac{1}{\lambda} - \frac{3}{\lambda}e^{-2} + \frac{2}{\lambda}e^{-2} = \left(1 - e^{-2}\right)\frac{1}{\lambda}$$

20. Markov Chains

Ky is stressed out by exams, and wants to use a Markov chain to improve their study habits. This chain has state space, $\mathscr{X} = \{0, 1, 2, 3\}$, which corresponds to what they are doing for one hour, where state 0 means "Locking In," state 1 means "Eating," state 2 means "Chilling," and state 3 means "Sleeping." The transition probabilities for this system are:

$$P = \begin{bmatrix} 1/2 & 1/4 & 1/4 & 0 \\ 1/2 & 1/4 & 1/4 & 0 \\ 1/2 & 0 & 1/4 & 1/4 \\ 0 & 1/4 & 0 & 3/4 \end{bmatrix}$$

1. This Markov chain is irreducible.

   Answer: True

2. What is the period of this Markov chain?

   Answer: 1 (there are self-loops).

3. Justify that the probability converges to a stationary distribution.

   Answer: The Fundamental Theorem of Markov Chains says that every irreducible Markov chain with period 1 converges to a unique stationary distribution.

4. (4 points) Let $\pi = [\pi(0)\ \pi(1)\ \pi(2)\ \pi(3)]$ be the stationary distribution. Give a set of equations that could be solved to find the stationary distribution.

   Answer:

$$\begin{aligned} \pi(0) &= (1/2)\pi(0) + (1/2)\pi(1) + (1/2)\pi(2) \\ \pi(1) &= (1/4)\pi(0) + (1/4)\pi(1) + (1/4)\pi(3) \\ \pi(2) &= (1/4)\pi(0) + (1/4)\pi(1) + (1/4)\pi(2) \\ \pi(3) &= (1/4)\pi(2) + (3/4)\pi(3) \end{aligned}$$

   You also need the fact that the probilities sum to 1:

$$\pi(0) + \pi(1) + \pi(2) + \pi(3) = 1\ .$$

5. (5 points) If Ky is currently locking in, what is the expected number of hours before they sleep?

   Answer: 16.

   For this we need the first-step equations for hitting time ($\beta(i)$ is the expected time of first visit to state 3 (sleeping) when starting from state i). Here are the equations:

$$\begin{aligned} \beta(0) &= 1 + (1/2)\beta(0) + (1/4)\beta(1) + (1/4)\beta(2) \\ \beta(1) &= 1 + (1/2)\beta(0) + (1/4)\beta(1) + (1/4)\beta(2) \\ \beta(2) &= 1 + (1/2)\beta(0) + (1/4)\beta(2) + (1/4)\beta(3) \\ \beta(3) &= 0 \end{aligned}$$

   Solving these gives $\beta(0) = 16$, $\beta(1) = 16$, and $\beta(2) = 12$. "Locking in" is state 0, so the answer is $\beta(0) = 16$.