CS 70 Discrete Mathematics and Probability Theory
Summer 2025 Tate

# Final

PRINT Your Name: ______________________, ______________________
(last) (first)

PRINT Your Student ID: ______________________

PRINT Your Exam Room: ______________________

SID of the person sitting to your left: ______________________

SID of the person sitting to your right: ______________________

SID of the person sitting in front of you: ______________________

SID of the person sitting behind you: ______________________

**Read This.**

- 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.
- Due to the above. Please move on. There are lots of problems to get points from. Do not get stuck. This is good advice anyway. In fact, we repeat it below.
- **Anything written outside the boxes provided will not be graded.**

**Advice.**

- 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. All blanks are worth 3 points each unless otherwise specified. 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.**
- The question statement is your friend. Reading it carefully is a tool to get to your "rational place".
- You may consult only *two 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.
- **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.**
- There are a total of 328 points on this exam, with 20 top-level questions.

**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?

**3. Propositional Logic**

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

   ○ True ○ False

   .

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

   ○ True ○ False

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

   ○ True ○ False

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

   ○ True ○ False

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

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

**4. Basic Proof Techniques**

1. The following predicate is part of the basis for a specific proof technique – what is it?
   $\bigwedge_{n=1}^{k} P(n) \implies P(k+1)$

2. (6 points) Let $x, y \in \mathbb{R}$. Prove that if $\frac{x+y}{2} = a$, then either $x \leq a$ or $y \leq a$.

**5. Graphs**

1. Any graph in which every vertex has degree at most 2 is planar.

   ◯ True ◯ False

2. Any graph that has a vertex of degree 6 is not planar.

   ◯ True ◯ False

3. A simple cycle with $n$ vertices is bipartite iff $n$ is odd.

   ◯ True ◯ False

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?

   (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.]

   (c) (4 points) Prove that the 4-dimensional hypercube is non-planar.

## 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.

   ◯ True ◯ False

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

   ◯ True ◯ False

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

   ◯ True ◯ False

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.

   ◯ True ◯ False

## 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.

   ◯ True ◯ False

   (b) (4 points) Give a one sentence justification of your answer.

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.

   ◯ True ◯ False

   (b) (4 points) Give a one sentence justification of your answer.

3. Prove the following two claims.

   (a) (6 points) Prove that for any modulus $m$, $x^2 \equiv (m-x)^2 \pmod{m}$.

   (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 $(n+1)$.

**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 $(x, y)$?

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?
   (b) If Kyle received a share of the form $(3, y)$, what is $y$?

3. A message is sent using Reed-Solomon coding with polynomial $P(x)$ over $GF(11)$, transmitting packets $P(1), P(2), \dots, P(8)$
   (a) If the code can correct for up to 2 corrupted packets, what is the length of the message (in packets)?
   (b) If a message is received and the Berlekamp-Welch algorithm computes polynomial $E(x) = x^2 - 6x + 8$, which packets are corrupted?

## 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?

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?

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?

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?

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?

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?

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?

**10. Countability and Computability**

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

   ◯ True ◯ False

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

   ◯ True ◯ False

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.

   ◯ True ◯ False

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.

   ◯ True ◯ False

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.

   ◯ True ◯ False

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.

## 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 $\mathbb{P}[A]$?

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

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

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

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

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

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?

**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 $p_1 + p_2 = 1$?

   ○ Always True ○ Always False ○ Depends

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?

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

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?

**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?

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?

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?

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 $\mathbb{P}(N_i)$?

   (b) Are the $N_i$ events pairwise independent?

   ◯ Yes ◯ No

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

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

**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:

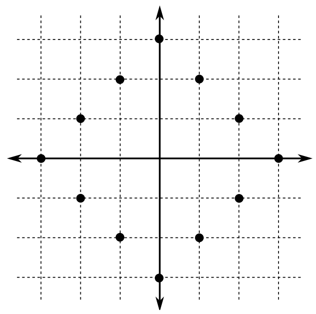
    > Given $X_1, X_2, \ldots, X_n$, we define $Y = X_1 + X_2 + \cdots + X_{n-1}$, and apply the induction hypothesis to get $\mathrm{Var}(Y) = \mathrm{Var}(X_1 + X_2 + \cdots + X_{n-1}) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_{n-1})$. Now we can apply the lemma to get
    > $$\mathrm{Var}(X_1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

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

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.

**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?

$\mathbb{E}[X]$: ☐  $\mathbb{E}[Y]$: ☐  $\mathrm{Var}[Y]$: ☐  $\mathrm{cov}(X,Y)$: ☐

(b) Are $X$ and $Y$ independent?

○ Yes ○ No

(c) Give a brief (but mathematically precise) justification for your previous answer.

(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?

$\mathbb{E}[X_1]$: ☐  $\mathbb{E}[Y_1]$: ☐  $\mathrm{Var}[Y_1]$: ☐  $\mathrm{cov}(X_1,Y_1)$: ☐

(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?

$\mathbb{E}[X_2]$: ☐  $\mathbb{E}[Y_2]$: ☐  $\mathrm{Var}[Y_2]$: ☐  $\mathrm{cov}(X_2,Y_2)$: ☐

**16. Distributions and Bounds**

1. Let $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ be independent.
   (a) Find $\mathbb{P}[X+Y=k]$.

   (b) What is distribution of $X+Y$?

   (c) (5 points) Give an intuitive interpretation for why $X+Y$ has the distribution you identified in part (b), without using formulas for probabilities.

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 $X \geq 90$.

   (b) For this distribution, $\mathbb{P}[X \leq 10] = \mathbb{P}[X \geq 90]$.

   ○ True ○ False

   (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).

**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?

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?

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 $c$?

   $c$: Expected winnings:

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 $c$?

   $c$: Expected winnings:

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?

**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 $R$?

   (b) What is the expected value for $R$?

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

   (d) Evaluate the integral from the previous part to give $\mathbb{P}[T_1 \leq T_2]$.

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 $0 \leq x \leq 10$.

   (b) What is the expected value for $S$?

   (c) (5 points) What is the probability that the two packets arrive within one time unit of each other (in other words, $\mathbb{P}[|T_1 - T_2| \leq 1]$)?

**19. Apply the Clamps**

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

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

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):

**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.

   ◯ True ◯ False

2. What is the period of this Markov chain?

3. Justify that the probability converges to a 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.

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