CS 70 Discrete Mathematics and Probability Theory
Fall 2023 Tal and Rao

Final

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

PRINT Your Student ID: ______________________________________

PRINT Your Exam Room: ______________________________________

SID of person to left: ____________________ SID of person to right: ____________________

SID of person in front: ____________________ SID of person in back: ____________________

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.**
- You may consult only *three 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 can leave your answers in terms of mathematical expressions such as e^{-10} or $\ln(2)$, but you may not leave your answer as a summation or integral (unless otherwise specified).
- **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 295 points on this exam, with 17 total 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 instructor and GSIs, 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 have any other browsers open while 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.
- I won't work for people from Stanford. [This is optional.]

Signed: ______________________________

2. Propositions. Basics.

1. Let P, Q and R be propositions.
   (a) $P \wedge \neg P \equiv$ True.
   ◯ True ◯ False
   (b) Consider the statement: $(P \implies R) \wedge (P \implies \neg R) \equiv$ ______. (Fill in the blank to make it true.)
   (c) $(P \wedge Q \implies \neg R) \equiv (\neg R \vee \neg P \vee \neg Q)$.
   ◯ True ◯ False
2. Let $P(x,y)$, $Q(x,y)$ be predicates on x and y.
   (a) $\neg(\forall x)(\exists y)(Q(x,y)) \equiv (\exists x)(\forall y)(\neg Q(x,y))$.
   ◯ True ◯ False
   (b) If $(\forall x)(\exists y)(Q(x,y)) \implies (\forall x)(\forall y)(P(x,y))$, and $P(a,b)$ is false for some a and b, then $(\exists x)$ such that $(\forall y)$, $Q(x,y)$ is ______.
   ◯ True ◯ False

3. Quick Proofs!

Prove or give a counterexample.

1. (5 points) For $a, b, c \in \mathbb{N}$, if $a \mid b$ and $b \mid c$, then $a \mid c$.

2. (7 points) Show that if the sum of digits of $n \in \mathbb{N}$ is divisible by 9, then n is divisible by 9.

4. Stable Matchings.

1. If a job is paired to its favorite candidate (the candidate that is first in its preference list) in some stable pairing, it must be matched to its favorite candidate in every stable pairing.

   ◯ True ◯ False

2. Consider a job that is last on every candidate's preference list, then it must be rejected $n-1$ times in the job-propose matching algorithm (for an n job, n candidate instance.)

   ◯ True ◯ False

3. If a job proposes to its least preferred candidate in the job-proposes matching algorithm, then the algorithm finishes on that day.

   ◯ True ◯ False

5. Graphs.

1. Consider G_e to be result of adding an edge e to a graph G. Then G_e either contains a cycle containing e or it has fewer connected components than G.

   ◯ True ◯ False

2. Consider a tree T with $n > 0$ vertices.

   (a) How many edges does T have?

   (b) Removing a subset of k edges from T forms a graph with _______ connected components on the vertices of T.

   (c) If T has exactly 2 vertices of degree 5 then it must have at least ______ vertices of degree one. (Give a tight bound.)

3. Consider removing a simple cycle from an n vertex Eulerian graph (graph that contains an Eulerian tour) resulting in a graph with c components.
   (a) Each connected component of the resulting graph is Eulerian.

   ◯ True ◯ False

   (b) The length of the simple cycle is at least ______. (Perhaps in terms of c and n; give a tight bound.)

4. A hypercube of dimension 4 has an Eulerian tour.

   ◯ True ◯ False

5. Given a graph $G = (V, E)$ that can be vertex-colored with k colors and $G' = (V, E')$ that is bipartite, then $G'' = (V, E \cup E')$ can be vertex-colored with ______ colors. (Give a tight bound.)

6. Consider a connected simple planar graph with n vertices.
   (a) If it has $2n$ edges, what is the number of faces in terms of n?

   (b) Given that the average face size is 5, what is the number of edges in terms of n?

6. Modular Arithmetic

1. Consider the equation $47(a)+49(b)=1$, for integers a and b. Find a solution of the form (a,b) that satisfies the above equation.

2. What is $2^{30} \pmod{35}$?

3. Let p, q be distinct primes.

   (a) Simplify $q^p \pmod{pq}$. (Hint: use CRT.)

   (b) There are $pq - (\_\_\_\_\_\_)$ elements of $\{0,\ldots,pq-1\}$ that have inverses under arithmetic modulo pq.

7. Mod: Proof.

1. (5 points) Prove that for a prime p, any $a \in \{2, \ldots, p-2\}$ satisfies $a^{-1} \not\equiv a \pmod{p}$. (Think about the equation $x^2 \equiv 1 \pmod{p}$)

2. What is the product of all elements of $\{1, \ldots, p-1\}$ modulo p? (Simplify.)

3. (5 points) For distinct primes $p, q > 2$, show that there exists an $a \in \{2, \ldots, pq-2\}$ where $a^{-1} \equiv a \pmod{pq}$. (Hint: perhaps use CRT.)

8. Polynomials.

1. Give a polynomial of degree 2 modulo 5, with roots at $x = 1$ and $x = 0$.

2. Let $P(x)$ be a polynomial of degree exactly d_P and let $Q(x)$ be a distinct polynomial of degree exactly d_Q. What is the maximum number of distinct values of x that satisfy $P(x) = Q(x)$?

3. Rao constructed a degree 1 polynomial $P(x) = mx + b \pmod{p}$ where p is prime, where b is his bank account password, and m is a random value. Rao has three kids; he gives the oldest kid $P(1)$, gives the second child $P(2)$, and gives the youngest the coefficient m.

   (a) Any two kids can find Rao's password.

   ◯ True ◯ False

   (b) Give an expression for finding the password b, given m and $P(1)$. (Answer should be possibly in terms of m, $P(1)$ and/or their inverses or constants modulo p.)

9. Countability/Computability

1. The set of subsets of the prime numbers is countable.

   ◯ True ◯ False

2. Given a set of sets, $\mathscr{A} = \{A_1, A_2, \ldots, A_n\}$, and the set $D = \{i \mid i \notin A_i\}$, then $D \notin \mathscr{A}$.

   ◯ True ◯ False

3. The set of programs that halt on the input 1 is countable.

   ◯ True ◯ False

4. There is a program that takes a program P, an input x, and a natural number n, and determines whether $P(x)$ runs for more than n steps.

   ◯ True ◯ False

10. Counting.

1. How many distinct ways can you rearrange the letters in the word "CAT"?

2. A histogram with k bins and n observations assigns a nonnegative integer to each bin, such that the values add up to n. How many such histograms are there? (Your answer should be in terms of n and k.)

3. Consider a staircase of height n which consists of k steps, where each step is of height 1 or 2. (Answers should be in terms of n and k.)

   (a) How many steps of height 2 are there exactly?

   (b) How many height n staircases are there with k steps?

11. Discrete Probability, Conditional Probability, Independence

1. There are two bags of marbles. The first bag has 4 red marbles and 4 blue marbles. The second bag has 3 red marbles and 5 blue marbles. You pick one bag uniformly at random and draw 2 marbles from it without replacement. Let E be the event that both marbles you drew were blue. Let A be the event that you picked the first bag.
   (a) Compute $\mathbb{P}[E \mid A]$.
   (b) Compute $\mathbb{P}[A \mid E]$.
2. You toss two fair dice (with labels 1 through 6).
   (a) Let X be their sum modulo 3. What is the probability that $X = 1$?
   (b) Let Y be their sum modulo 4. What is the probability that $Y = 3$?
   (c) What is $\mathbb{P}[(X = 1) \cap (Y = 3)]$?
   (d) What is $\mathbb{P}[(X = 1) \cup (Y = 3)]$?

3. You play "Let's Make a Deal" but the host, Monty Hall, doesn't know where the prize is. The setting is as follows:

   - Carol picks where to put the prize uniformly at random from $\{1,2,3\}$, but tells no one (including Monty).
   - You first pick door number 1.
   - Monty picks one of the other doors $\{2,3\}$ at random and opens it. In the case he opens the door with the prize, you automatically lose. Otherwise, he offers to let you switch to the other closed door.

   Let A be the event that Monty opened a door *without* the prize. Let B be the event that door number 1 has the prize.

   (a) Compute $\mathbb{P}[A]$.

   (b) Compute $\mathbb{P}[B \mid A]$.

12. Discrete Random Variables

1. We throw n distinguishable balls into n distinguishable bins. Let X be the number of balls in bin 1.

   (a) Compute $\mathbb{P}[X=0]$.

   (b) What is the distribution of X?

   (c) As n grows to infinity, what does $\mathbb{P}[X=10]$ converge to? Your answer should not contain a limit.

2. We flip a fair coin 100 times. Let X be the number of heads minus the number of tails.

   (a) Compute $\mathbb{P}[X=0]$.

   (b) What is $\mathbb{E}[X]$?

   (c) (4 points) Let $Y=X^2$. Are X and Y independent? (Justify briefly)

   (d) What is $\mathrm{Cov}(X,Y)$?

13. Expectation, Variance, Covariance

1. Gavin has n different pairs of socks (n left socks and n right socks, for $2n$ individual socks total) and is doing his laundry. He notices that the laundry machine spits out a uniformly random permutation of the $2n$ socks.

   Let X be the number of matching pairs that are placed next to each other.

   As an example, for the outcome $1_L 3_R 1_R 2_R 2_L 3_L$, $X = 1$ since only the 2nd pair of socks are placed next to each other.

   (a) What is the probability that the 1st pair of socks are placed next to each other?

   (b) What is $\mathbb{E}[X]$?

   (c) What is the probability that both the 1st pair are placed next to each other and the 2nd pair are placed next to each other?

   (d) Let p be the answer to (a), and let q be the answer to (c). What is $\mathrm{Var}(X)$? Feel free to leave your answer in terms of p and q.

2. In a class with k students, each student has a uniformly random birthday in $\{1, \dots, 365\}$, and the students' birthdays are independent.

   (a) What is the expected number of pairs of students that share a birthday?

   (b) What is the expected number of triplets of students that share a birthday?

   (c) What is the variance of the number of pairs of students that share a birthday?

14. Concentration Inequalities, Confidence Intervals, LLN

We flip a biased coin with unknown heads probability p, and want to estimate $\frac{1}{p}$. Let X be the number of flips until we get k heads.

1. (5 points) Show that $Y = \frac{X}{k}$ is an unbiased estimator for $\frac{1}{p}$; that is, $\mathbb{E}[Y] = \frac{1}{p}$.

2. Compute $\text{Var}(Y)$.

3. Use Chebyshev's inequality to upper bound $\mathbb{P}\left[\left|Y - \frac{1}{p}\right| \geq \varepsilon \cdot \frac{1}{p}\right]$.

4. Suppose p is unknown, and we only know $p \geq 0.5$. What value of k should we pick to ensure $\mathbb{P}\left[|Y - \frac{1}{p}| < 0.1 \cdot \frac{1}{p}\right] \geq 0.95$?

15. Continuous Probability

1. (10 points) You drive from Berkeley to Los Angeles. The distance is 360 miles. You choose a speed (in miles per hour) uniformly at random from the interval $[40,60]$ and drive. What is the expected driving time in hours? (Show your work.)

2. Let $X,Y,Z \sim U([0,1])$ be mutually independent.
   (a) What is the CDF $F_1(t)$ of $\max(X,Y,Z)$, for $t \in [0,1]$?

   (b) What is the PDF $f_1(t)$ of $\max(X,Y,Z)$, for $t \in [0,1]$?

   (c) What is the expectation of $\max(X,Y,Z)$?

   (d) What is the PDF $f_2(t)$ of $\min(X,Y,Z)$, for $t \in [0,1]$?

   (e) The median of three numbers is the middle number after sorting, e.g., $\text{median}(0.6,0.3,0.9) = 0.6$. What is the expectation of $\text{median}(X,Y,Z)$?

3. Let $T_S \sim N(4, 0.2)$ and $T_J \sim N(4, 0.3)$ be two independent random variables.

   (a) What's the distribution of $T_S + T_J$?

   (b) What's the distribution of $T_S - T_J$?

   (c) Compute $\mathbb{P}[T_S > T_J]$.

16. Conditional Expectation/Estimation

1. Suppose a store has 3 circular dartboards with radii a, b and c. Shreyas is in a rush so he buys one dartboard at random and leaves the store. Suppose his throws are uniformly distributed across the entire dartboard. Shreyas throws a dart at random at his board.

   (a) What is the dart's expected distance to the center? Your answer may depend on the variables a, b, and c.

   (b) (5 points) Justify your answer.

2. Suppose we have a random variable Y with expectation 7 and variance 2. We also have $Z \sim N(3,5)$, where Y and Z are independent. Let $X = Y + Z$.

   (a) (3 points each) Given that $Y = y$, what is the function $\hat{X}$ in terms of y that minimizes the mean squared error $\mathbb{E}\left[(X-\hat{X})^2 \mid Y = y\right]$? What is the corresponding mean squared error?

   $\hat{X} =$ ☐ Mean Squared Error $=$ ☐

   (b) What is the linear least squares estimator $L[X \mid Y]$?

   ☐

3. (7 points) Suppose the height of a person (in inches) is distributed like $X+Y$ where X and Y are independent; $X \sim N(69,3)$ is a genetic component, and $Y \sim N(0,1)$ is an environmental component.

   Two identical twins are separated at birth and raised in different environments.

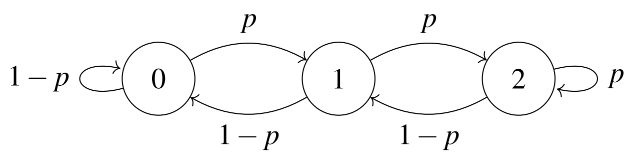
   The random variables $Z_1 = X + Y_1$ and $Z_2 = X + Y_2$ are the heights of twin A and twin B respectively (note that they share the genetic component). Here, $X \sim N(69,3)$, $Y_1 \sim N(0,1)$, and $Y_2 \sim N(0,1)$ are mutually independent.

   What is the linear least squares estimator of Z_2 given Z_1? (i.e., $L[Z_2 \mid Z_1]$)

   ☐

17. Markov Chains

1. Consider the following Markov chain:



Let X_n be the random variable representing the current state of the Markov chain at step n.

(a) (5 points) Let $p = \frac{1}{3}$. Write, but do not solve, a system of linear equations to find the stationary distribution of this Markov chain.

Suppose we choose a p such that the stationary distribution of the Markov chain is

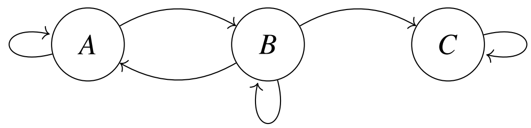
$$\pi = \begin{bmatrix} \frac{1}{13} & \frac{3}{13} & \frac{9}{13} \end{bmatrix}.$$

Let Y_n be a new random variable dependent on X_n such that $Y_n = 0$ if $X_n \in \{0, 1\}$, and $Y_n = 1$ if $X_n = 2$. Suppose that in this experiment we only observe the value of Y_n, and we'd like to estimate the value of X_n given our observations. Further, suppose that in our observations, the Markov chain has already been evolving over a long period of time, and X_n has converged to the stationary distribution.

(b) Compute the minimum mean squared estimate of X_n given Y_n, i.e. $\mathbb{E}[X_n \mid Y_n]$.

(c) Compute the linear least squares estimate of X_n given Y_n, i.e. $L[X_n \mid Y_n]$. Your answer should be in the form $aY_n + b$, for some constants a and b. (Hint: there should not be much work involved.)

2. Suppose you have a fair die (with labels 1 through 6) and you keep rolling the die until the product of the last two rolls is 15. Consider the following Markov chain modeling this scenario.



Here, A is the starting state, as well as the state in which the previous roll was in $\{1,2,4,6\}$, B is the state in which the previous roll was in $\{3,5\}$ but the product of the last two rolls is not 15, and C is the state in which the product of the last two rolls was 15.

(a) (1 point each) What are the transition probabilities in this Markov chain? You only need to provide the transition probabilities of the drawn edges in the diagram above.

$\mathbb{P}[X_n = A \mid X_{n-1} = A] =$ ☐ $\mathbb{P}[X_n = B \mid X_{n-1} = B] =$ ☐

$\mathbb{P}[X_n = B \mid X_{n-1} = A] =$ ☐ $\mathbb{P}[X_n = C \mid X_{n-1} = B] =$ ☐

$\mathbb{P}[X_n = A \mid X_{n-1} = B] =$ ☐ $\mathbb{P}[X_n = C \mid X_{n-1} = C] =$ ☐

(b) This Markov chain is irreducible.

○ True ○ False

(c) This Markov chain is aperiodic.

○ True ○ False

(d) (6 points) Let $\beta(i)$ denote the expected number of rolls you make until the last two rolls multiply to 15, given that you are currently in state i of the Markov chain. Write, but do not solve, a system of linear equations involving $\beta(i)$ for $i \in \{A,B,C\}$. (You should use the variables in the following diagram as the coefficients in your equations.)