Discrete Mathematics and Probability Theory
Ayazifar and Rao

Final

PRINT Your Name: ______________________________, ______________________________
(Last) (First)

READ AND SIGN The Honor Code:
As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others.
Signed: ______________________________

PRINT Your Student ID: ______________________________

WRITE your exam room: ______________________________

WRITE the name of the person sitting to your left: ______________________________

WRITE the name of the person sitting to your right: ______________________________

PLEASE READ THE FOLLOWING INSTRUCTIONS CAREFULLY

- After the exam starts, please *write your student ID on every page*. **You will not be allowed to write** *anything* **once the exam ends**.
- We will not grade anything outside of the space provided for a problem unless we are clearly told in the space provided for the question to look elsewhere. We will not grade scratch paper, all work must be on exam.
- The questions vary in difficulty. If you get stuck on any one, it helps to leave it and try another one.
- In general, no justification on short answer/true false questions is required unless otherwise indicated. Write your answers in boxes where provided.
- Calculators are not allowed. **You do NOT need to simplify any probability related answers to a decimal fraction,** but your answer must be in the simplest form (no summations or integrals).
- You may consult only *3 sheets of notes*. Apart from that, you are not allowed to look at books, notes, etc. Any electronic devices such as phones and computers are NOT permitted.
- Regrades will be due quickly so watch piazza.
- There are **19** double sided pages on the exam. Notify a proctor immediately if a page is missing.
- You have **180** minutes: there are **6** sections with a total of **68** parts on this exam worth a total of **243** points.

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

1. Discrete Math: True/False (12 parts: 3 points each.)

1. $\forall x, \forall y, \neg P(x,y) \equiv \neg \exists y, \exists x, P(x,y)$

   ◯ True
   ◯ False

2. $(P \implies Q) \equiv (Q \implies P)$.

   ◯ True
   ◯ False

3. Any simple graph with n vertices can be colored with $n-1$ colors.

   ◯ True
   ◯ False

4. The set of all finite, undirected graphs is countable.

   ◯ True
   ◯ False

5. The function $f(x) = ax \pmod N$ is a bijection from and to $\{0,\ldots,N-1\}$ if and only if $\gcd(a,N) = 1$.

   ◯ True
   ◯ False

6. For a prime p, the function $f(x) = x^d \pmod p$ is a bijection from and to $\{0,\ldots,p-1\}$ when $\gcd(d,p-1) = 1$.

   ◯ True
   ◯ False

7. A male optimal pairing cannot be female optimal.

   ◯ True
   ◯ False

8. For any undirected graph, the number of odd-degree vertices is odd.

   ◯ True
   ◯ False

9. For every real number x, there is a program that given k, will print out the kth digit of x.

   ◯ True
   ◯ False

10. There is a program that, given another program P, will determine if P halts when given no input.

    ◯ True
    ◯ False

11. Any connected simple graph with n vertices and *exactly* n edges is planar.

    ◯ True
    ◯ False

12. Given two numbers, x and y, that are relatively prime to N, the product xy is relatively prime to N.

    ◯ True
    ◯ False

2. Discrete Math:Short Answer (10 parts: 4 points each)

1. If $\gcd(x,y) = d$, what is the least common multiple of x and y (smallest natural number n where both $x|n$ and $y|n$)? [Leave your answer in terms of x,y,d]

2. Consider the graph with vertices $\{0,\ldots,N-1\}$ and edges $(i, i+a) \pmod N$ for some $a \not\equiv 0 \pmod N$. Let $d = gcd(a,N)$. What is the length of the longest cycle in this graph in terms of some subset of N, a, and d?

3. What is the minimum number of women who get their favorite partner (first in their preference list) in a female optimal stable pairing? (Note that the minimum is over any instance.)

4. What is the number of ways to split 7 dollars among Alice, Bob and Eve? (Each person should get an whole number of dollars.)

5. What is $6^{24} \pmod{35}$?

6. If one has three distinct degree at most d polynomials, $P(x), Q(x), R(x)$, what is the maximum number of intersections across all **pairs** of polynomials?

   Recall that we define intersections to be two polynomials having the same value at a point. (That is if $P(1) = Q(1)$, and $P(2) = R(2)$ and $R(3) = Q(3)$, that is three intersections. If they all meet at a point $P(1) = Q(1) = R(1)$, that is three intersections.)

7. Working modulo a prime $p > d$, given a degree exactly d polynomial $P(x)$, how many polynomials $Q(x)$ of degree at most d are there such that $P(x)$ and $Q(x)$ intersect at exactly d points?

8. Recall that the vertices in a d-dimensional hypercube correspond to $0-1$ strings of length d. We call the number of 1's in this representation the **weight** of a vertex.

   (a) How many vertices in a d-dimensional hypercube have weight k?

   (b) How many edges are between vertices with weight at most k and vertices with weight greater than k?

9. How many elements of $\{0, \ldots, p^k - 1\}$ are relatively prime to p?

3. Some proofs. (3 parts. 5/5/8 points.)

1. Recall for x, y, with $gcd(x, y) = d$, that there are $a, b \in Z$ where $ax + by = d$. Prove that $gcd(a, b) = 1$.

2. You have n coins. The probability of the ith coin being heads is $1/(i+1)$ (i.e., the biases of the coins are $\frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n+1}$). You flip all the coins. What is the probability that you see an even number of heads? Prove it. (Hint: the answer is quite simple.)

3. Consider a game with two players alternating turns. The game begins with $N > 0$ flags. On each turn, each player can remove $1, 2, 3$, or 4 flags. A player wins if they remove the last flag (even if they removed several in that turn).

   Show that if both players play optimally, player 2 wins if N is a multiple of 5, and player 1 wins otherwise.

4. Probability:True/False. (7 parts, 3 points each.)

1. For a random variable X, the event "$X = 3$" is independent of the event "$X = 4$".

   ◯ True

   ◯ False

2. Let X, Y be Normal with mean μ and variance σ^2, independent of each other. Let $Z = 2X + 3Y$. Then, $LLSE[Z \mid X] = MMSE[Z \mid X]$.

   ◯ True

   ◯ False

3. Any irreducible Markov chain where one state has a self loop is aperiodic.

   ◯ True

   ◯ False

4. Given a Markov Chain, let the random variables $X_1, X_2, X_3, \ldots$, where $X_t =$ the state visited at time t in the Markov Chain. Then $E[X_t | X_{t-1} = x] = E[X_t | X_{t-1} = x \cap X_{t-2} = x']$.

   ◯ True

   ◯ False

5. Given an expected value μ, a variance $\sigma^2 \geq 0$, and a probability p, it is always possible to choose a and b such that a discrete random variable X which is a with probability p and b with probability $1 - p$ will have the specified expected value and variance.

   ◯ True

   ◯ False

6. Consider two random variables, X and Y, with joint density function $f(x, y) = 4xy$ when $x, y \in [0, 1]$ and 0 elsewhere. X and Y are independent.

   ◯ True

   ◯ False

7. Suppose every state in a Markov chain has exactly one outgoing transition. There is one state, s, whose outgoing transition is a self-loop. All other states' outgoing transitions are not self-loops. If a unique stationary distribution exists, it must have probability 1 on s and 0 everywhere else.

   ◯ True

   ◯ False

5. Probability: Short Answer. (17 parts, 4 points each.)

1. Consider $X \sim G(p)$, a geometric random variable X with parameter p. What is $Pr[X > i | X > j]$ for $i \geq j$?

2. Suppose we have a random variable, X, with pdf

$$f(x) = \begin{cases} cx^2, \text{if } 0 \leq x \leq 1 \\ 0, \text{ otherwise} \end{cases}$$

   What is c?

3. Given a binomial random variable X with parameters n and p, ($X \sim B(n, p)$) what is $Pr[X = E[X]]$? (You should assume pn is an integer.)

4. $Pr[A|B] = 1/2$, and $Pr[B] = 1/2$, and A and B are independent events. What is $Pr[A]$?

5. Aaron is teaching section and has 6 problems on the worksheet. The time it takes for him to finish covering each question are i.i.d. random variables that follow the exponential distribution with parameter $\lambda = 1/20$. Additionally, for each question, Aaron may choose to skip it entirely with probability $p = 1/3$. What is the expected time of section?

6. Let X be a uniformly distributed variable on the interval $[3, 6]$. What is Var(X)?

7. Label N teams as team 1 through team N. They play a tournament and get ranked from rank 1 to rank N (with no ties). All rankings are equally likely.

   (a) What is the total number of rankings where team 1 is ranked higher than team 2?

   (b) What is the expected number of teams with a strictly lower rank number than their team number? For example, if team 3 was rank 1, their rank number (1) is lower than their team number (3). Simplify your answer (i.e. no summations).

8. Let X be a random variable that is never smaller than -1 and has expectation 5. Give a non-trivial upper bound on the probability that X is at least 12.

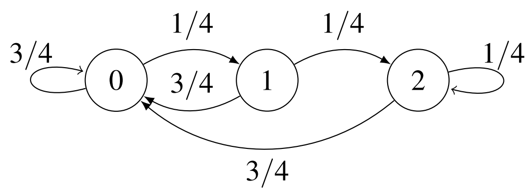
9. Let X be a random variable with mean $E[X] = 5$ with $E[X^2] = 29$. Give a non-trivial upper bound on the probability that X is larger than 12.

10. Let T be the event that an individual gets a positive result on a medical test for a disease and D be the event that an individual has the disease. The test has the property that $Pr[T|D] = .9$ and $Pr[T|\overline{D}] = .01$. Morever, $Pr[D] = .01$. Given a positive result, what the probability that the individual has a disease? (No need to simplify your answer, though it should be a complete expression with numbers.)

11. Let R be a continuous random variable corresponding to a reading on a medical test for an individual and D be the event that the individual has a disease. The probability of an individual having the disease is p. Further, let $f_{R|D}(r)$ (and $f_{R|\overline{D}}(r)$) be the conditional probability density for R conditioned on D (respectively conditioned on $\overline{D}$). Given a reading of r, give an expression for the probability the individual has the disease in terms of $f_{R|D}(r), f_{R|\overline{D}}(r)$, and p.

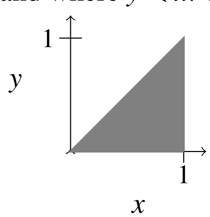
12. For continuous random variables, X and Y where $Y = g(X)$ for some differentiable, bijective function $g : \mathbb{R} \to \mathbb{R}$. What is $f_Y(y)$ in terms of $f_X(\cdot)$, $g(\cdot)$, $g^{-1}(\cdot)$ and $g'(\cdot)$? (Possibly useful to remember that $f_Y(y)dy = Pr[y \le Y \le y + dy]$.)

13. What is the stationary distribution, π, for the following three state Markov chain? (Hint: $\pi(0) = 3/4$)



$\pi(1)$ $\pi(2)$

14. Consider continuous random variables, X and Y, with joint density that is $f(x,y) = 2$ for $x, y \in [0,1]$ and where $y < x$. That is, the distribution is uniform over the shaded region in the figure below.



Say someone takes a sample of X or Y with equal probability, and then announces that the value is $2/3$. What is the probability that the sample is from X?

15. Given a random variable $X \sim \text{Expo}(\lambda)$, consider the integer valued random variable $K = \lceil X \rceil$.

(a) What is $Pr[K = k]$?

(b) What standard distribution with associated parameter(s) does this correspond to?

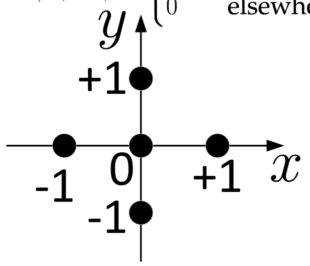
6. Longer Probability Questions.

1. **[I iterated my expectations, and you can, too!] (4 parts. 5 points each.)**

Consider two discrete random variables X and Y. For notational purposes, X has probability mass function (or distribution), $p_X(x) = Pr[X = x]$, mean μ_X, and variance σ_X^2. Similiarly, random variable Y has PMF $p_Y(y) = Pr[Y = y]$, mean μ_Y and variance σ_Y^2.

For each of True/False parts in this problem, either prove the corresponding statement is True in general or use exactly one of the counterexamples provided below to show the statement is False.

$$p_{X,Y}(x,y) = \begin{cases} 1/5 & \text{At each } \bullet \\ 0 & \text{elsewhere.} \end{cases}$$



(a) Potential Counterexample I

The PMF for random variable X is

$$p_X(x) = \mathsf{Pr}(X = x) = \begin{cases} 1/3 & x = -1, 0, +1 \\ 0 & \text{elsewhere.} \end{cases}$$

Random Variable Y is

$$Y = X^2 \quad \text{for all X.}$$

(b) Potential Counterexample II

(a) Suppose $E[Y|X] = c$, where c is a fixed constant. This means that the conditional mean $E[Y|X]$ does *not* depend on X.

i. Show that $c = \mu_Y$, the mean of Y.

ii. True or False?

The random variables X and Y are independent.

iii. True or False?

The random variables X and Y are *uncorrelated*, meaning that $cov(X,Y) = 0$.

(b) Suppose X and Y are *uncorrelated*, meaning that $cov(X,Y) = 0$.
True or False?

The conditional mean is $E[Y|X] = c$, where c is a fixed constant, meaning that $E[Y|X]$ does *not* depend on X.

2. **[Estimations of a random variable with noise.] (6 parts. 2/4/2/2/4/8 points.)**

   Let random variable Y denote the blood pressure of a patient, and suppose we model it as a Gaussian random variable having mean μ_Y and variance σ_Y^2.

   Our blood pressure monitor (measuring device) is faulty. It yields a measurement

$$X = Y + W$$

   where the noise W is a zero mean Gaussian random variable ($\mu_W = 0$) with variance σ_W^2. Assume that the noise W is *uncorrelated* with Y. Note, that the actual blood pressure Y is inaccessible to us, due to the additive noise W.

   (a) Show that $\sigma_X^2 = \sigma_Y^2 + \sigma_W^2$.

   (b) Show that $L(Y|X)$, the Linear Least-Square Error Estimate for the blood pressure Y, based on the measured quantity X, is given by

$$L(Y|X) = a + bX, \quad \text{where} \quad a = \frac{\sigma_W^2}{\sigma_Y^2 + \sigma_W^2}\,\mu_Y \quad \text{and} \qquad b = \frac{\sigma_Y^2}{\sigma_Y^2 + \sigma_W^2}.$$

(c) We now consider two extreme cases.

i. Suppose the blood pressure monitor has been repaired —that is, it introduces no noise. Determine a simple expression for $L(Y|X)$ in this case.

ii. Suppose the blood pressure monitor's performance has deteriorated, so it now introduces noise whose variance $\sigma_W^2 \gg \sigma_Y^2$. In the limit $\sigma_W^2 \to \infty$, what does your best linear estimator converge to? Explain briefly, in plain English words, why your answer makes sense.

(d) Recall $L[Y|X]$ is a function of X and is a random variable. Let $\hat{Y} = L[Y|X] = a + bX$. Determine the distribution of $\widehat{Y}$ and the appropriate parameters.

(e) We estimate $\hat{\mu}_Y$ of the true mean μ_Y as

$$\widehat{\mu}_Y = \frac{X_1 + \cdots + X_n}{n},$$

where X_i are independent measurements of the random variable $X = Y + W$.
We want to be *at least* 95% confident that the absolute error $|\widehat{\mu}_Y - \mu_Y|$ is within 4% of μ_Y. Your task is to determine the *minimum* number of measurements n needed so that

$$Pr\big[\,|\widehat{\mu}_Y - \mu_Y| \leq 0.04\,\mu_Y\big] \geq 0.95.$$

You may assume that $\sigma_Y^2 = 12$ and $\sigma_W^2 = 4$ and that the true mean $\mu_Y \in [60, 90]$.
(Remember that in this course, you may assume that a Gaussian random variable lies within 2σ of its mean with 95% probability.)

3. **[Derive the Unexpected from a Uniform PDF] (2 parts. 3/2 points.)**

   You wish to use $X \sim U[0,1)$ to produce a different *nonnegative* random variable $Y = -\frac{1}{\lambda}\ln(1-X)$, for $0 \leq X < 1$, where λ is a positive constant, and ln is the natural logarithm function.
   (Note that the pdf for $X \sim U[0,1)$ is the same as for $X \sim U[0,1]$.)

   (a) Determine the CDF $F_Y(y) = Pr[Y \leq y]$. [It may be useful to recall that $F_x(x) = x$ for $x \in [0,1)$.]

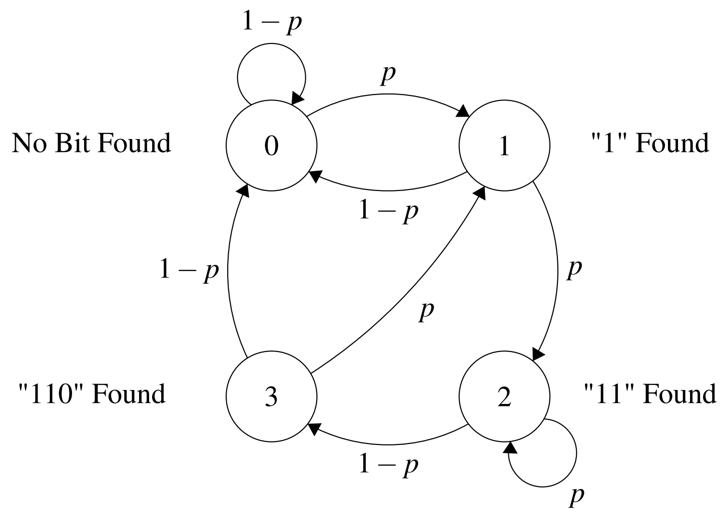
   (b) Determine the PDF $f_Y(y)$ and indicate what standard distribution it corresponds to.

4. **[Finding a Three-Bit String in a Binary Bitsream] (3 parts. 2/5/5 points.)**

Consider a bitstream $B_1, B_2, \ldots$ consisting of IID Bernoulli random variables obeying the probabilities $Pr[B_n = 1] = p$, and $Pr[B_n = 0] = 1 - p$, for every $n = 1, 2, \ldots$.

Here, $0 < p < 1$.

We begin parsing the bitstream from the beginning, in search of a desired binary string represented by the codeword $c = (1,1,0)$. We say that we've encountered the codeword c at time n if $(B_{n-2}, B_{n-1}, B_n) = (1,1,0)$. We model this process using the Markov chain shown below.



There are four states, labeled 0,1,2, and 3. The state number i represents the number of the leading (leftmost) bits of the codeword $c = 110$ for which we've found a match at time n—starting from the leading (leftmost) bit. For example, being in state 2, means you saw a 11 in the two latest bits.

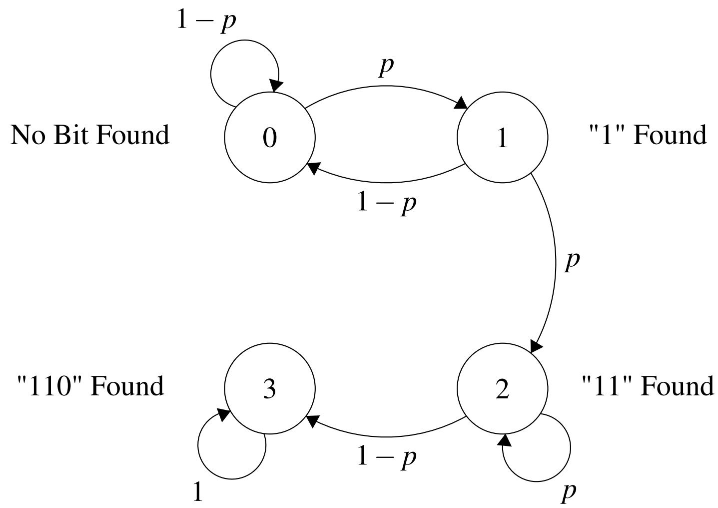
That is, if X_n denote the state of the process at time n and and the bit-stream consists of $B_1, \ldots, B_n$. We have $X_n = 2$ when $(B_{n-1}, B_n) = 11$. We begin with X_0 in state 0 by default which corresponds to no prefix of the codeword $c = 110$ has been read.

(a) Provide a clear, succinct explanation as to why the Markov chain above has a set of unique limiting-state (i.e., stationary) probabilities:

$$\pi_i = \lim_{n \to \infty} Pr[X_n = i], \qquad i = 0, 1, 2, 3.$$

(b) Determine a simple expression for the limiting-state probability π_3 of State 3.
To receive full credit, you must explain your answer.
Depending on how you tackle this part, you may need only a small fraction of the space given to you below.

(c) For the remainder of this problem, we want to find the *expected time* $E(N)$ until the first occurrence of the string $c = 110$ in the bitstream.
Accordingly, we remove all the outgoing edges from State 3 in the original Markov chain, and turn State 3 into an absorbing state having a self-loop probability of 1 as below.



Determine $E(N)$, the expected time at which we first enter State 3—that is, the time at which the string $c = (1,1,0)$ occurs for the first time.
Hint: We recommend that you break down N into two parts. Let $N = N_{02} + N_{23}$, where N_{02} denotes the number of steps until first passage into State 2, starting from State 0, and N_{23} denotes the number of steps it takes to transition for the first time from State 2 to State 3. Show that

$$E(N_{02}) = \frac{1}{p} + \frac{1}{p^2},$$

determine $E(N_{23})$, and put your results together to obtain $E(N)$.