CS70 Discrete Mathematics and Probability Theory, Fall 2018

Final Exam

8:00-11:00am, 14 December

Your First Name: **Your Last Name:**

SIGN Your Name: **Your SID Number:**

Your Exam Room:

Name of Person Sitting on Your Left:

Name of Person Sitting on Your Right:

Name of Person Sitting in Front of You:

Name of Person Sitting Behind You:

Instructions:

(a) *As soon as the exam starts, please write your student ID in the space provided at the top of every page! (We will remove the staple when scanning your exam.)*

(b) *There are 9* **double-sided** *sheets (18 numbered pages) on the exam. Notify a proctor immediately if a sheet is missing.*

(c) *We will not grade anything outside of the space provided for a question (i.e., either a designated box if it is provided, or otherwise the white space immediately below the question).* **Be sure to write your full answer in the box or space provided!** *Scratch paper is provided on request; however, please bear in mind that nothing you write on scratch paper will be graded!*

(d) *The questions vary in difficulty, so if you get stuck on any question it may help to leave it and return to it later.*

(e) *On questions 1-2: You must give the answer in the format requested (e.g., True/False, an expression, a statement.) An expression may simply be a number or an expression with a relevant variable in it. For short answer questions, correct, clearly identified answers will receive full credit with no justification. Incorrect answers may receive partial credit.*

(f) *On questions 3-8, you should give arguments, proofs or clear descriptions if requested. If there is a box you must use it for your answer.*

(g) *You may consult three two-sided "cheat sheets" of notes. Apart from that, you may not look at any other materials. Calculators, phones, computers, and other electronic devices are NOT permitted.*

(h) *You may, without proof, use theorems and lemmas that were proven in the notes and/or in lecture.*

(i) *You have 3 hours: there are 8 questions on this exam worth a total of 190 points.*

[Exam starts on next page]

1. **True/False** [*No justification; answer by shading the correct bubble. Points per answer as indicated; total of 39 points. No penalty for incorrect answers.*]

(a) Zero or more of the following are valid logical equivalences, for arbitrary propositions P and Q. Indicate which by shading the appropriate circles.

YES NO

○ ○ $(P \Rightarrow \neg Q) \equiv (Q \Rightarrow \neg P)$ *1pt*

○ ○ $(P \Rightarrow Q) \equiv (P \vee \neg Q)$ *1pt*

○ ○ $\neg(\neg P \vee Q) \equiv (P \wedge \neg Q)$ *1pt*

○ ○ $((P \Rightarrow Q) \wedge (Q \Rightarrow P)) \equiv ((P \wedge \neg Q) \vee (\neg P \wedge Q))$ *1pt*

○ ○ $\neg(P \vee (\neg P \wedge Q)) \equiv (\neg P \wedge \neg Q)$ *1pt*

(b) Define the following predicates involving the variables x, y over the universe of cats.

- $G(x)$: x has green eyes
- $B(x)$: x has a bushy tail
- $F(x, y)$: x is fatter than y

Consider the following statement:

"Every cat with green eyes is fatter than at least one cat that has a bushy tail but doesn't have green eyes."

Which (if any) of the following expressions are accurate translations of this statement? Answer by shading either the "Yes" or the "No" bubble for each expression. (There may be more than one "Yes" answer.)

YES NO

○ ○ $\forall x \exists y (G(x) \wedge B(y) \wedge \neg G(y) \wedge F(x, y))$ *1pt*

○ ○ $\exists y \forall x (\neg G(y) \wedge B(y) \wedge F(y, x))$ *1pt*

○ ○ $\forall x \forall y (G(x) \Rightarrow (B(y) \wedge \neg G(y) \wedge F(x, y)))$ *1pt*

○ ○ $\forall x \exists y (G(x) \Rightarrow (B(y) \wedge \neg G(y) \wedge F(x, y)))$ *1pt*

[Q1 continued on next page]

(c) Consider the following stable marriage instance, consisting of four men 1, 2, 3, 4 and four women A, B, C, D:

Man	Women			
1	A	B	D	C
2	D	C	B	A
3	A	C	D	B
4	B	A	C	D

Woman	Men			
A	2	3	1	4
B	2	3	1	4
C	1	4	2	3
D	4	2	1	3

For each of the following statements about this instance, indicate whether the statement is True or False by shading the corresponding bubble.

TRUE FALSE

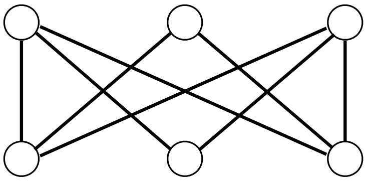
○ ○ The pairing (1,D), (2, C), (3, A), (4, B) is stable. *2pts*

○ ○ The pairing (1, C), (2, B), (3, A), (4, D) is male pessimal. *2pts*

○ ○ There exists a stable pairing in which man 4 is paired with woman A. *2pts*

○ ○ Woman A is paired with man 3 in every stable pairing. *2pts*

(d) Consider the following graph.



Which of the following properties is/are true of this graph?

TRUE FALSE

○ ○ The graph is connected. *1pt*

○ ○ The graph is bipartite. *1pt*

○ ○ The graph is planar. *1pt*

○ ○ The graph has a Hamiltonian cycle. *1pt*

[Q1 continued on next page]

(e) A connected graph G (with no self-loops or multiple edges between the same pair of vertices) has seven vertices whose degrees are 3, 4, 4, 4, 5, 6, 6 respectively. Answer each of the following by shading the appropriate bubble: “Yes”, “No” or “??” if the answer cannot be determined from the given information.

YES	**NO**	**??**		
○	○	○	Does G have an Eulerian tour?	*1pt*
○	○	○	Does G have an Eulerian walk?	*1pt*
○	○	○	Is G planar?	*1pt*

(f) Classify each of the following functions $f : \mathbb{R} \to \mathbb{R}$ as (i) neither 1-1 nor onto; (ii) 1-1 but not onto; (iii) onto but not 1-1; (iv) both 1-1 and onto (a bijection).

(i)	**(ii)**	**(iii)**	**(iv)**		
○	○	○	○	$f(x) = x^2$.	*1pt*
○	○	○	○	$f(x) = x + 1$.	*1pt*
○	○	○	○	$f(x) = \frac{1}{x}$ for $x \neq 0$, $f(0) = 0$.	*1pt*
○	○	○	○	$f(x) = e^x$.	*1pt*

(g) Answer each of the following questions **TRUE** or **FALSE** by shading the appropriate bubble.

TRUE	**FALSE**		
○	○	There are two distinct powers of 2 that are equal mod 97.	*1pt*
○	○	For sets A, B, if A is uncountable and B is countable, then the difference $A \setminus B$ is uncountable.	*1pt*
○	○	For a collection of countably infinite sets A_i, $i \in \mathbb{N}$, the union $\bigcup_{i \in \mathbb{N}} A_i$ is uncountable.	*1pt*
○	○	If A is uncountable and B is finite and non-empty, then $A \times B$ is countable.	*1pt*

[Q1 continued on next page]

(h) Answer each of the following questions **TRUE** or **FALSE** by shading the appropriate bubble.

TRUE	**FALSE**		
○	○	For all events $A_1, \ldots, A_n$, we have $\mathbb{P}[\bigcup_{i=1}^n A_i] \leq \sum_{i=1}^n \mathbb{P}[A_i]$.	*1pt*
○	○	For dependent random variables X, Y and constants a, b, it is possible that $\mathbb{E}[aX + bY] \neq a\mathbb{E}[X] + b\mathbb{E}[Y]$.	*1pt*
○	○	$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ if and only if X and Y are independent.	*1pt*
○	○	Consider two random variables X and Y with ranges $\mathcal{A}_X$ and $\mathcal{A}_Y$, respectively. If there exist $a \in \mathcal{A}_X$ and $b \in \mathcal{A}_Y$ such that $\mathbb{P}[X = a, Y = b] = \mathbb{P}[X = a]\mathbb{P}[Y = b]$, then X and Y are independent.	*1pt*
○	○	Consider the standard Coupon Collector's Problem with n coupon types, and let W_n denote the total number of trials required to collect all n coupon types. Then, $\lim_{n\to\infty} \mathbb{E}[W_n]/(n \ln n) = 1$.	*1pt*
○	○	Suppose $P(x) = Ax + B$ is a random polynomial where A and B are independent standard normal random variables. Then $P(x)$ has a root with probability 1.	*1pt*
○	○	For all Markov chains $\{X_n, n \in \mathbb{N}\}$ with finite state space S and $\forall m \in \{0, 1, \ldots, n\}$, $\mathbb{P}[X_n = j \mid X_0 = i] = \sum_{k \in S} \mathbb{P}[X_n = j \mid X_m = k] \times \mathbb{P}[X_m = k \mid X_0 = i]$.	*1pt*

2. **Short Answers** [*Answer is a single number or expression;* **write it in the box provided: anything outside the box will not be graded;** *no justification necessary. 3 points per answer; total of 78 points. No penalty for incorrect answers.*]

   (a) What is the inverse of the function $f(x) = 3x - 2$ over the reals? *3pts*

   (b) Compute $\gcd(323, 152)$. *3pts*

   (c) What is $37^{225} \bmod 113$? [Note: 113 is prime.] *3pts*

   (d) Solve for x in the modular equation $5x - 7 = 0 \pmod{11}$. *3pts*

   (e) Suppose Alice's public RSA key is $(N, e) = (143, 7)$. What is the value d that Alice needs in order to decrypt messages sent to her? (Hint: Note that $143 = 11 \times 13$.) *3pts*

   (f) Rex chooses a random polynomial P of degree at most k over GF(q) (for a prime $q > k$), by selecting $k + 1$ coefficients independently and uniformly at random from $\{0, 1, \dots, q - 1\}$.

      (i) What is the probability that Rex's polynomial goes through the point $(0, 0)$, i.e., that $P(0) = 0$? *3pts*

      (ii) What is the probability that Rex's polynomial has exactly k distinct roots? *3pts*

[Q2 continued on next page]

(g) Alice wants to share a secret (a number mod 7) among her four loyal companions in such a way that any three of the companions can recover the secret, but no two of them can. All of them have taken CS70 and agree to do polynomial secret sharing with the secret stored at $P(0)$ for a suitable polynomial P over $\mathrm{GF}(7)$. Alice gives each of her companions one of the following points: $(1,1),(2,0),(3,0),(4,1)$. What is the secret? *3pts*

(h) Laurel wants to send a message consisting of 100 packets to Hardy over a noisy channel. Laurel knows that up to $\frac{1}{6}$ of the packets sent may be corrupted by the channel. Assuming Laurel uses the Berlekamp-Welch encoding scheme, how many packets must he send to ensure that Hardy will be able to recover the original message? *3pts*

(i) We say that program P_1 *dominates* program P_2 if P_1 halts on every input on which P_2 halts. The problem *Dominates* takes as input two programs, P_1 and P_2, and decides whether P_1 dominates P_2. The following pseudo-code gives a reduction from the Halting Problem, *Halt*, to *Dominates*. Fill in the blank to make the reduction behave correctly. *3pts*

```
Test-Halt(P,x)
    let $P_1$ be a program that, on every input □
    let $P_2$ be a program that, on every input, halts
    if Test-Dominates($P_1$,$P_2$) then return "yes" else return "no"
```

(j) Suppose there are k keys $\{w_1,\ldots,w_k\}$ and a hash table of size n. How many distinct hash functions are there such that keys w_1 and w_2 do not get mapped to the same location of the hash table? *3pts*

(k) A 5-card poker hand is called a *straight* if its cards can be re-arranged to form a contiguous sequence, regardless of their suits, i.e., if the hand is of the form $\{A,2,3,4,5\},\{2,3,4,5,6\},\ldots,$ or $\{10,J,Q,K,A\}$. How many straight hands are there consisting of 3 black and 2 red cards? *3pts*

[Q2 continued on next page]

(l) Consider the complete graph K_n with $n \geq 3$ vertices, and suppose that each edge is colored blue with probability p and red with probability $1 - p$, independently of all other edges. Let Δ_n denote the number of completely blue triangles resulting from this random coloring.

(i) What is $\mathbb{E}[\Delta_n]$? *3pts*

(ii) Use the union bound to find a lower bound on $\mathbb{P}[\Delta_n = 0]$. *3pts*

(m) Let A and B denote two events such that $A \subset B$. Suppose $\mathbb{P}[A] = a$ and $\mathbb{P}[B] = b$, and let I_A and I_B denote the indicator random variables for A and B, respectively. Find $\text{Cov}(I_A, I_B)$. *3pts*

(n) Consider an urn with 3 blue balls and 1 red ball, and suppose you sample one ball at a time with replacement. Let X be the number of draws required until both of the colors, blue and red, have been observed at least once.

(i) What is $\mathbb{P}[X = n \mid \text{The first ball drawn is red}]$, for $n \geq 2$? *3pts*

(ii) What is $\mathbb{P}[X = n]$, for $n \geq 2$? *3pts*

[Q2 continued on next page]

(o) Suppose $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$ are independent random variables, where $\lambda, \mu > 0$. What is $\mathbb{P}[\min\{X, Y\} \leq t]$, where t is a positive constant? *3pts*

(p) Suppose $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$ are independent random variables, where $\lambda, \mu > 0$. What is the conditional probability $\mathbb{P}[X < Y \mid \min\{X, Y\} > t]$, where t is a positive constant? *3pts*

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

(r) Suppose A and B are independent $\text{Normal}(1, 1)$ random variables. Find $\mathbb{P}[2A + B \geq 4]$ in terms of the cumulative distribution function (c.d.f.) Φ of the standard normal distribution. *3pts*

(s) Consider randomly dropping a circular coin of radius 1 cm onto a large rectangular grid where horizontal lines are 3 cm apart, while vertical lines are 4 cm apart. What is the probability that the coin intersects at least one grid line? *3pts*

[Q2 continued on next page]

(t) Let X be a continuous random variable with probability density function (p.d.f.) $f(x) = 2x$ if $0 \leq x \leq 1$, and $f(x) = 0$ otherwise. Find Var$[X^2]$. *3pts*

(u) Find a and b such that the following function F is a valid c.d.f. for a continuous random variable, and find the corresponding p.d.f. $f(x)$: *3pts*

$$F(x) = \begin{cases} 0, & \text{for } x \leq 0, \\ a(1-x)^2 + b, & \text{for } 0 < x < 1, \\ 1, & \text{for } x \geq 1. \end{cases}$$

$a =$ □ , $b =$ □ , $f(x) =$ □ .

(v) Consider a two-state Markov chain with transition probability matrix $\boldsymbol{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$, where $0 < a, b < 1$. Find the corresponding stationary distribution $\boldsymbol{\pi} = (\pi_1, \pi_2)$. *3pts*

$\pi_1 =$ □ and $\pi_2 =$ □

(w) There are two candy jars labeled A and B. Each day, you choose jar A with probability p and jar B with probability $1 - p$, and eat one of the candies from the chosen jar. Let $\alpha(i, j)$ denote the probability that jar A becomes empty before jar B, given that jar A currently contains i candies while jar B contains j candies. For $i, j > 0$, write down a recursive formula satisfied by $\alpha(i, j)$: *2pts*

Write down the boundary conditions (base cases) for your recursion: *1pt*

3. Trees [*All parts to be justified. Total of 10 pts.*]

(a) Prove by strong induction on the number of vertices that the vertices of any tree can be colored with two colors so that no two adjacent vertices get the same color. [Hint: Remove an arbitrary edge from the tree.] *5pts*

(b) Let T be a tree with n vertices. For $1 \leq i \leq n-1$, let n_i denote the number of vertices in T of degree exactly i. Show that $\sum_{i=1}^{n-1} i n_i = 2(n-1)$. [Reminder: You may use without proof any results from notes or lecture, provided that they are clearly stated.] *2pts*

(c) Suppose $n \geq 6$ is even and suppose that T has $\frac{n}{2}+2$ vertices of degree 1. Prove using part (b) that T must have at least one vertex of degree at least 4. *3pts*

4. Wilson's Theorem [*All parts to be justified. Total of 12 pts.*]

This question leads you through a proof of Wilson's Theorem, which says the following:

Theorem: *A natural number* $n > 1$ *is prime* **if and only if** $(n-1)! \equiv -1 \pmod{n}$.

(a) Assume first that $n > 1$ is not prime, and let $1 < q < n$ be a divisor of n. Show that if a number $a \equiv b \pmod{n}$ then also $a \equiv b \pmod{q}$. *3pts*

(b) Deduce from the previous part that, when $n > 1$ is not prime, $(n-1)! \not\equiv -1 \pmod{n}$. *3pts*

(c) Now assume that $n > 1$ is prime. Let x be any number in the range $1 \leq x \leq n-1$. Show that the only such x which are their own inverse $\pmod{n}$ are $x = 1$ and $x = n-1 \equiv -1 \pmod{n}$. [Hint: Think about polynomials!] *3pts*

(d) Deduce from the previous part that, when $n > 1$ is prime, $(n-1)! \equiv -1 \pmod{n}$. *3pts*

5. Partitions via Random Sampling [*No justification necessary. Total of 13 points.*]

There are two urns: Urn 1 has 10 blue and 6 red marbles, while Urn 2 has 7 blue and 9 red marbles. Alice, Bob, and Carol decide to divide up the marbles among them using random sampling. They will first choose one of the urns uniformly at random (u.a.r.), and then use the following scheme to sample from the same chosen urn until it is empty: In each round of sampling, one person is chosen u.a.r., and that person will sample a marble u.a.r. from the urn and keep the marble. Let B_A, B_B, B_C respectively denote the number of blue marbles that Alice, Bob, and Carol have at the end; R_A, R_B, R_C are similarly defined for red marbles.

Whenever possible, express all combinatorial factors in terms of binomial coefficients.

(a) Find $\mathbb{P}[B_A = 3, B_B = 5, B_C = 2 \mid \text{Urn 1 was chosen}]$. *3pts*

(b) Find $\mathbb{P}[B_A = 4, R_A = 2 \mid \text{Urn 1 was chosen}]$. *3pts*

(c) Let E denote the event that 3 blue and 2 red marbles are observed in the first five rounds of sampling. Find $\mathbb{P}[E \mid \text{Urn 1 was chosen}]$ and $\mathbb{P}[\text{Urn 1 was chosen} \mid E]$. *4pts*

(d) What is the total number of distinct partitions of the marbles among Alice, Bob, and Carol, given that Urn 1 was chosen? [Remark: Marbles of the same color are indistinguishable, but marbles of different colors are distinguishable.] *3pts*

6. Probability Bounds and Limits [*No partial credit for (a)-(d). Total of 13 points.*]

Consider i.i.d. random variables $X_1, X_2, \ldots$ with probability distribution $\mathbb{P}[X_i = 2] = \frac{1}{4}$, $\mathbb{P}[X_i = 4] = \frac{1}{2}$, and $\mathbb{P}[X_i = 6] = \frac{1}{4}$ for all $i = 1, 2, \ldots$. Let $S_n = X_1 + X_2 + \cdots + X_n$.

(a) Find $\mathbb{E}[S_n]$ and $\mathrm{Var}[S_n]$. *3pts*

$\mathbb{E}[S_n] =$ □

$\mathrm{Var}[S_n] =$ □

(b) Markov's inequality implies which of the following? Shade only one bubble. *2pts*

○ $\mathbb{P}[S_n < 5n] \geq \frac{1}{5}$

○ $\mathbb{P}[S_n < 5n] \leq \frac{1}{5}$

○ $\mathbb{P}[S_n < 5n] \geq \frac{4}{5}$

○ $\mathbb{P}[S_n < 5n] \leq \frac{4}{5}$

○ None of the above.

(c) Chebyshev's inequality implies which of the following? Shade only one bubble. *3pts*

○ $\mathbb{P}[S_n < 5n] \geq \frac{1}{n}$

○ $\mathbb{P}[S_n < 5n] \leq \frac{1}{n}$

○ $\mathbb{P}[S_n < 5n] \geq 1 - \frac{1}{n}$

○ $\mathbb{P}[S_n < 5n] \leq 1 - \frac{1}{n}$

○ None of the above.

(d) Let Φ denote the c.d.f. of the standard normal distribution. For large n, which of the following is true? Shade only one bubble. *2pts*

○ $\mathbb{P}[S_n \geq 4n + \epsilon n] \approx \Phi(\sqrt{\frac{n}{2}}\,\epsilon)$

○ $\mathbb{P}[S_n \geq 4n + \epsilon n] \approx 1 - \Phi(\sqrt{\frac{n}{2}}\,\epsilon)$

○ $\mathbb{P}[S_n \geq 4n + \epsilon n] \approx \Phi(\frac{\epsilon}{2})$

○ $\mathbb{P}[S_n \geq 4n + \epsilon n] \approx 1 - \Phi(\frac{\epsilon}{2})$

○ None of the above.

(e) For $\delta > 0$, $\lim_{n\to\infty} \mathbb{P}[S_n \leq (4 - \delta)n] =$ □ . Justify your answer below. *3pts*

7. Poisson Distribution [*Justification required where stated. Total of 12 points.*]

Assume that the number of data blocks received at a data storage center per month follows a Poisson distribution with rate $\lambda > 0$, and assume that these numbers over different months are mutually independent. After each month of storage, each data block has probability $p > 0$ of getting corrupted, independently of all other data blocks. Let X_0 denote the number of new data blocks received this month, and, for $n \in \mathbb{Z}^+$, let X_n denote the number of data blocks received n months ago that have so far not been corrupted.

(a) Prove that $X_1 \sim \text{Poisson}[(1-p)\lambda]$. *5pts*

[Q7 continued on next page]

(b) For $n \in \mathbb{Z}^+$, what is the distribution of X_n? State its name and specify its parameter(s). No justification necessary. *2pts*

(c) What is the distribution of $X_0 + X_1 + \cdots + X_n$? State its name and specify its parameter(s). No justification necessary. *2pts*

(d) $\lim\limits_{\lambda \to \infty} \mathbb{P}\big[X_0 - \lambda < \sqrt{\lambda}\big] = \boxed{\quad\quad}$.

[Your answer may be left as an unevaluated sum or integral.] Justify your answer below. *3pts*

8. I.I.D. Continuous Uniform Random Variables [*All parts to be justified. Total of 13 points.*]

For $n \geq 2$, let $X_1, \ldots, X_n$ be *independent* Uniform$[0, 1]$ random variables, and, for $i \in \{1, \ldots, n\}$, let Y_i denote the ith smallest value of $\{X_1, \ldots, X_n\}$. For example, $Y_1 = \min\{X_1, \ldots, X_n\}$, while $Y_n = \max\{X_1, \ldots, X_n\}$. In HW 12, you found the distributions of Y_1 and Y_n.

(a) Prove that the probability density function (p.d.f.) of Y_2 is given by $f_{Y_2}(y) = n(n-1)y(1-y)^{n-2}$, where $0 \leq y \leq 1$. [Hint: First derive the cumulative distribution function of Y_2.] *4pts*

(b) For the case of $n = 2$, find the joint p.d.f. $f(y_1, y_2)$ of Y_1 and Y_2. Justify your answer. *3pts*

[Q8 continued on next page]

(c) Assume again $n = 2$ and let $G = Y_2 - Y_1$, the gap size between Y_1 and Y_2. Find the p.d.f. of G. Justify your answer. *3pts*

(d) What is $\mathbb{P}[G > \frac{1}{2}]$? Justify your answer. [Hint: You should be able to solve this problem without using the p.d.f. of G.] *3pts*

[End of Exam]