Final Exam

Last name	First name	SID

Name of student on your left:
Name of student on your right:

- DO NOT open the exam until instructed to do so.
- The total number of points is **110**, but a score of $\geq$ **100** is considered perfect.
- You have 10 minutes to read this exam without writing anything and 150 minutes to work on the problems.
- Box your final answers.
- **Remember to write your name and SID on the top left corner of every sheet of paper.**
- **Do not write on the reverse sides of the pages.**
- All eletronic devices must be turned off. Textbooks, computers, calculators, etc. are prohibited.
- No form of collaboration between students is allowed. If you are caught cheating, you may fail the course and face disciplinary consequences.
- **You must include explanations to receive credit.**

Problem	Max	Points	Problem	Max	Points
1	12		7	10	
2	16		8	12	
3	8		9	12	
4	8		10	8	
5	8		11	1	
6	15				
Total				110	

Problem 1. (12 pts) You must give brief explanations in the provided boxes to get any credit.

(a) Let $\hat{X}$ be the cubic (functions of the form $f(z) = az^3 + bz^2 + cz + d$) least squares estimate of X given the observation Y. True or False: $\text{cov}(X - \hat{X}, Y^2) = 0$.

True or False:
Explanation:

(b) Let X and Y be two Jointly Gaussian random variables such that $\text{cov}(X, Y) = 0$. True or False: X and Y need not be independent.

True or False:
Explanation:

(c) True or False: $\text{var}(X) = E[\text{Var}(X|Y)] + \text{Var}(E[X|Y])$ for random variables X and Y.

True or False:
Explanation:

(d) Explain in the box below why, in general, the Chernoff bound is a "better" bound than Chebyshev and why Chebyshev is a "better" bound than Markov.

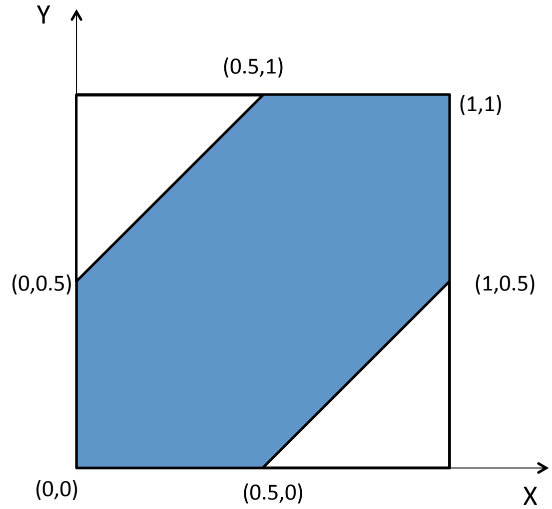
Explanation:

Problem 2. (16 pts) Parts $(a), (b)$ and (c) are short answer questions and unrelated to each other.

(a) (4 pts) Consider an Erdos-Renyi random graph $G = G(n, p)$. Let the random variable X give the number of *isolated* nodes in the graph. Your friend Bollobas claims that $\mathrm{Var}(X) = nq(1-q)$, where q is the probability a node is isolated. Is Bollobas correct? You must explain your answer to receive any credit.

(b) (6 pts) HKN wants to take a survey of how many Berkeley EECS undergraduate students actually cheat on their exams. As this is a sensitive question and the HKN officers know that almost nobody would admit to cheating even if they did, they devise a privacy-preserving scheme as follows: Each survey-taking student privately flips a fair coin. If the coin comes up heads, he/she answers the question truthfully. If it comes up tails, he/she answers the question randomly, i.e. equally likely to be "yes" and "no". Each student cheats on his or her exam with probability p, and each student acts independently. If 200 students respond to the survey such that there are 70 respondents who said "yes" and 130 who said "no", what is the maximum-likelihood estimate (MLE) of p?

(c) (6 pts) Consider the following figure:



Suppose that (X, Y) are uniformly distributed on the shaded region. Find the linear least squares estimate of X given Y $(L[X|Y])$.

Problem 3. (8 pts) Consider IID random variables $X_i \sim N(0,1)$. Find the minimum mean square estimate (MMSE) of $X_1 + X_2 + X_3$ given the observations $X_1 + X_2$, $X_2 + X_3$, and $X_3 + X_4$.

Problem 4. (8 pts) Suppose X, Y, Z are mutually independent random variables, each of which is uniformly distributed on $[0, 1]$. Find the minimum mean square estimate (MMSE) of $(X + Y)^2$ given the observation $Y + Z$.

Problem 5. (8 pts) You observe one realization of a random variable Y and would like to determine which distribution it came from. Consider the hypotheses:

$$X = 0 : Y \sim N(0, 1)$$
$$X = 1 : Y \sim N(0, 2)$$

Design a hypothesis test that maximizes the probability of correct detection ($P(\hat{X} = 1 | X = 1)$) and has the property that the probability of false alarm $P(\hat{X} = 1 | X = 0) \leq 0.05$.

Problem 6. (15 pts) Reddit is running an experiment known as "the button". The button has a t second timer. Everytime it is pressed, its clock resets to t seconds. If the clock expires, the button shuts down forever. You may assume that the button was pressed at time 0.

(a) (10 pts) Users arrive to the webpage according to a Poisson process with rate λ. Upon arriving, they decide to press the button with probability p, otherwise they leave. In addition to the users, there is an administrator for Reddit online 24 hours a day who refreshes the page and presses the button according to an independent Poisson process with rate 2λ. What is the expected amount of time until the button expires?

(b) (5 pts) For this part, assume no time limit on the button (that is $t \to \infty$). Suppose that you observe the third time the button was pressed and the first time the button was pressed. Find the linear least square estimate (LLSE) of the time at which the button was pressed for the second time.

Problem 7. (10 pts) A particle is moving randomly according to the following update equations: at time $i = 1, 2, \ldots$, the particle's position is given by:

$$X(i) = 2X(i-1) + V(i),$$

with initial position $X(0)$. At each time $j \geq 0$, you have access only to a noisy measurement of the position:

$$Y(j) = X(j) + W(j).$$

Assume that $X(0), V(i), W(i)$ are IID $N(0,1)$ for all $i \geq 0$.

(a) (4 pts) Find $\hat{X}(0) = E[X(0)|Y(0)]$

(b) (6 pts) Your friend Rudolf wants to estimate $X(1)$ from the new observation $Y(1)$ that comes in, and the estimate $\hat{X}(0)$ from part (a). In other words, he wants to form the estimate $X\hat{(}1) = E[X(1)|Y(0), Y(1)]$ in the following form:

$$\hat{X}(1) = \alpha\hat{X}(0) + \beta Y(1).$$

Find α and β and use a diagram to help explain your reasoning.

Problem 8. (12 pts) Consider an infinite continuous-time Markov chain X_t with $t \geq 0$, depicted in Fig. 1. Let $\mathbf{X_0} = (\mathbf{0}, \mathbf{2})$ and T be the first time the chain visits state $(0,0)$. Find $E[T]$ in terms of λ, μ and α. You will receive partial credit for setting up a correct set of equations.

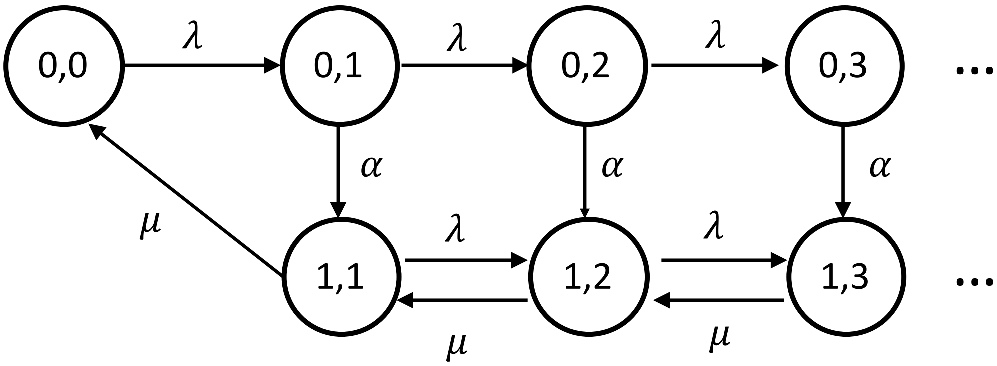
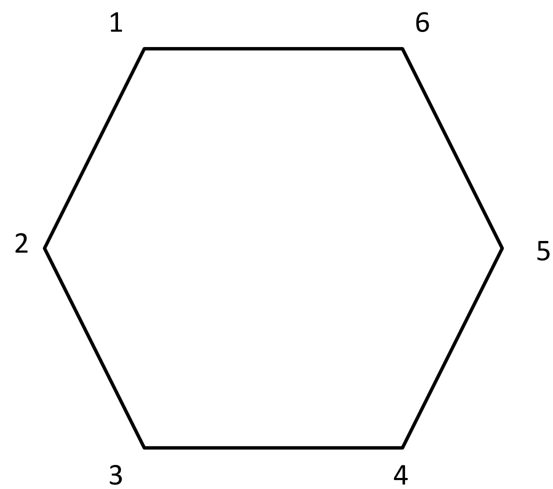


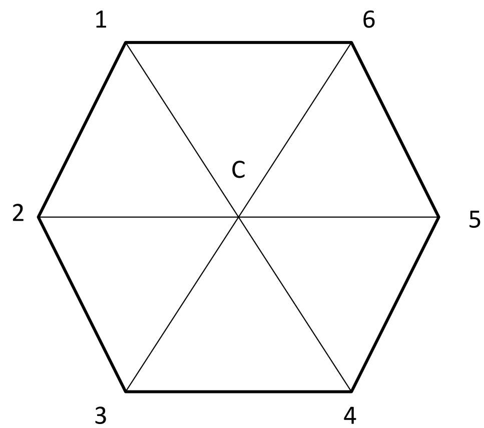
Figure 1: Infinite continuous-time Markov chain

Problem 9. (12 pts) An ant is performing a random walk on the vertices $\{1, 2, 3, 4, 5, 6\}$ of the hexagon below: at each time, it randomly selects one of its neighbors, with each neighbor equally likely to be selected, and moves to that neightbor. Note that for both parts (a) and (b), you will receive partial credit for setting up a correct set of equations.



(a) (6 pts) If the ant starts at vertex 1, what is the expected amount of time until it returns to vertex 1?

(b) (6 pts) An exterminator puts diagonal lines on the hexagon and places a trap at point C. Supposing that the ant performs a random walk on the vertices $\{1, 2, 3, 4, 5, 6, C\}$ and starts the walk at point 1, what is the probability that the ant returns to point 1 before hitting point C?



Problem 10. (8 pts) The final exam scores of 10 students taking a graduate EECS class are: $34, 45, 50, 56, 60, 74, 80, 81, 95, 100$. Suppose everyone gets either an A or a B, with the scores of students receiving A grades being distributed as a Normal with mean μ_A and variance σ^2, and the scores of students receiving B grades as Normal with mean μ_B and variance σ^2. Use a Hard-EM algorithm to cluster the scores into A and B grade clusters and we initialize the algorithm with $\mu_A = 80$ and $\mu_B = 50$. What will be estimates of μ_A and μ_B produced once the algorithm converges?

Problem 11. (1 point) Please leave any feedback for the course staff here. What did you like and dislike about the course? What can we improve upon?

END OF THE EXAM.

Please check whether you have written your name and SID on every page.

Hope you enjoyed the class! You learned a lot!