

Name (last, first): _____

Student ID: _____

Write your name and PID on the top of **EVERY PAGE**.

Write the solutions to each problem on separate pages. **CLEARLY INDICATE** on the top of each page the number of the corresponding problem. Different parts of the same problem can be written on the same page (for example, part (a) and part (b)).

The exam consists of 4 questions. Your answers must be carefully justified to receive credit.

This exam will be scanned. Make sure you write **ALL SOLUTIONS** on the paper provided. **DO NOT REMOVE ANY OF THE PAGES**.

No calculators, phones, or other electronic devices are allowed.

Remember this exam is graded by a human being. Write your solutions **NEATLY AND COHERENTLY**, or they risk not receiving full credit.

You are allowed to use one 8.5 by 11 inch sheet of paper with handwritten notes (on both sides); no other notes (or books) are allowed.

This exam is property of the regents of the university of California and not meant for outside distribution. If you see this exam appearing elsewhere, please **NOTIFY** the instructor at ynemish@ucsd.edu and the UCSD Office of Academic Integrity at aio@ucsd.edu.

1. (20 points) Let X be a random variable taking values in the set $\{1, 2, 3, \dots\}$. Let $p = P(X = 1)$ satisfy $0 < p < 1$.

Suppose that for random variable X

$$P(X = k + n | X > n) = P(X = k) \quad (1)$$

for any $n, k \geq 1$.

- (a) Consider the identity (1) with $k = 1$ and $n = 1$

$$P(X = 2 | X > 1) = P(X = 1).$$

Use it to express $P(X = 2)$ in terms of p . [Hint: Notice that $\{X = 2\} \subset \{X > 1\}$.]

Solution.

Using (1) for $k = 1$ and $n = 1$ and the definition of the conditional probability we have

$$p = P(X = 2 | X > 1) = \frac{P(X = 2, X > 1)}{P(X > 1)} = \frac{P(X = 2)}{1 - P(X = 1)} = \frac{P(X = 2)}{1 - p}, \quad (2)$$

from which we find that

$$P(X = 2) = (1 - p)p. \quad (3)$$

- (b) Consider the identity (1) with $k = 2$ and $n = 1$

$$P(X = 3 | X > 1) = P(X = 2).$$

Use it together with the result of (a) to express $P(X = 3)$ in terms of p .

Solution. Using (1) for $k = 2$ and $n = 1$, equation (3) and the definition of the conditional probability we have

$$(1 - p)p \stackrel{(3)}{=} P(X = 2) \stackrel{(1)}{=} P(X = 3 | X > 1) = \frac{P(X = 3, X > 1)}{P(X > 1)} = \frac{P(X = 3)}{1 - p}, \quad (4)$$

from which we find that

$$P(X = 3) = (1 - p)^2 p. \quad (5)$$

(CONTINUED ON THE NEXT PAGE)

(c) Use the identity (1) with general $k \geq 1$ and $n = 1$ to show that

$$P(X = k + 1) = P(X = k)P(X > 1),$$

and determine the distribution of X .

Solution. Repeating the same argument as in parts (a) and (b) for general $k \geq 1$ and $n = 1$ we get

$$P(X = k) \stackrel{(1)}{=} P(X = k + 1 | X > 1) = \frac{P(X = k + 1, X > 1)}{P(X > 1)} = \frac{P(X = k + 1)}{1 - p}, \quad (6)$$

from which it follows that

$$P(X = k + 1) = (1 - p)P(X = k). \quad (7)$$

Fix $l \geq 2$. By applying (7) $l - 1$ times we find

$$P(X = l) = (1 - p)P(X = l - 1) \quad (8)$$

$$= (1 - p)^2 P(X = l - 2) \quad (9)$$

$$= \dots \quad (10)$$

$$= (1 - p)^{l-1} P(X = 1) \quad (11)$$

$$= (1 - p)^{l-1} p. \quad (12)$$

We conclude that X has geometric distribution with parameter $p = P(X = 1)$, $X \sim \text{Geom}(p)$.

Remark. Property (1) is the discrete version of the memoryless property. If $X \sim \text{Geom}(p)$, then clearly X satisfies the memoryless property (1). In this problem we have proven the converse: if X is a discrete random variable satisfying (1), then $X \sim \text{Geom}(p)$. Therefore, geometric distribution is *the only* discrete distribution on the set $\{1, 2, \dots\}$ that satisfies the memoryless property (1).

2. (20 points) Let $X \sim \text{Poisson}(\lambda)$. Compute

$$E\left(\frac{1}{1+X}\right).$$

Solution. This is a direct calculation:

$$\begin{aligned} E\left(\frac{1}{1+X}\right) &= \sum_{k=0}^{\infty} \left(\frac{1}{1+k}\right) P(X = k) \\ &= \sum_{k=0}^{\infty} \left(\frac{1}{1+k}\right) \left(e^{-\lambda} \frac{\lambda^k}{k!}\right) \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{1}{1+k} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{(k+1)!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{\lambda(k+1)!} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{(k+1)!} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \\ &= \frac{e^{-\lambda}}{\lambda} \left(\left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right) - 1 \right) \\ &= \frac{e^{-\lambda}}{\lambda} (e^{\lambda} - 1) \\ &= \frac{1 - e^{-\lambda}}{\lambda}, \end{aligned}$$

though you did not need to simplify to the last line. The penultimate line would also receive full credit.

3. (20 points) Suppose that we plan to interview n randomly chosen individuals to estimate the unknown fraction $p \in (0, 1)$ of the population that likes ice cream. Let $\hat{p} = \frac{S_n}{n}$ be the random variable that records the proportion of the individuals who say they do like ice cream. How many people must we interview to have at least a 95% chance of capturing the true fraction p with a margin of error .01? You may leave your answer in terms of the inverse Φ^{-1} of the CDF of the standard normal.

Solution. We start with the formula

$$2\Phi(2\varepsilon\sqrt{n}) - 1 \geq .95.$$

We are told $\varepsilon = .01$. So, this simplifies to

$$\Phi(.02\sqrt{n}) \geq .975.$$

Since the density φ of the standard normal is positive everywhere, its CDF is strictly increasing and hence invertible. This allows us to conclude that

$$.02\sqrt{n} \geq \Phi^{-1}(.975).$$

In other words, we need

$$n \geq (50\Phi^{-1}(.975))^2 = 2500[\Phi^{-1}(.975)]^2.$$

4. (20 points) Show that there is no random variable X such that

$$E(e^X) = 3 \quad \text{and} \quad E(e^{2X}) = 4. \tag{13}$$

Solution. Suppose that there exists a random variable X for which (13) holds. Consider the random variable $Y = e^X$. Then we have that

$$E(Y) = 3, \quad E(Y^2) = 4. \tag{14}$$

From this we find that the variance of Y is

$$\text{Var}(Y) = E(Y^2) - E(Y)^2 = 4 - 3^2 = -5 < 0, \tag{15}$$

which is impossible.

