CS 177 Midterm Exam 1

Applications of Probability in Computer Science, Winter 2024

Thursday, February 8, 2024 from 3:30-4:50pm

Your Name: **Row/Seat Number:**

Your ID #(e.g., 123456789) **UCINetID (e.g., ucinetid@uci.edu)**

- Total time is **75 minutes**. Please look through all questions and organize your time wisely.
- Please put your name and ID **on every page**.
- Please write your **row and seat number** on the cover page.
- **Electronic devices are not allowed.**
- For full credit, be sure to **write clearly** and **show all your work**.
- You may use **one** (two-sided) 8.5x11-inch sheet of (your own) handwritten notes.
  You may keep these notes, they do not need to be turned in.
- Have a question? **Please raise your hand for help.**

Problems

Total: *(60 points)*

This page is intentionally blank, use as you wish.

Name: ________ ID#: ________

Problem 1 Early Disease Detection. *(15 points)*

Researchers are investigating the spread of an infectious disease and the number of symptoms shown by infected individuals. The number of major symptoms X can be 0, 1 or 2. The disease test result Y is binary, where 1 indicates a positive test and 0 a negative test. The following table defines the joint probability distribution of X and Y:

x	y	$p_{XY}(x, y)$
0	0	0.40
0	1	0.05
1	0	0.10
1	1	0.15
2	0	0.05
2	1	0.25

When computing probabilities or expected values below, **enter the numerical value of your answer in the provided box**, and also include equations showing how you calculated that answer.

a) *What is the probability that a patient will test positive for the disease?*

$$P(Y = 1) = \boxed{}$$

b) *If a patient shows at least one of these symptom, what is the probability that they are nevertheless healthy? To determine thus, compute the following conditional probability:*

$$P(Y = 0 \mid X > 0) = \boxed{}$$

c) What is the expected number of symptoms?

$$E[X] = \boxed{}$$

d) Are the random variables X and Y are independent? Justify your answer.

Name: ______ ID#: ______

Problem 2 The Weather Forecast. *(20 points)*

In the mythical Land of the Anteaters, every day's weather is independent of every other day. Each morning, it starts raining with probability 0.2. No weather forecasts can exist, because every day's weather is completely unpredictable.

a) *What is the probability that during an entire 7-day week, it does not rain at all? (Your answer may involve an exponent, and does not need to be evaluated numerically.)*

$P(\text{Dry Week}) =$ ______

b) *You keep a journal of the weather for 30 days, and count the number of days X that it rains. (Note that $0 \le X \le 30$.) What is the probability mass function $p_X(x)$? Justify your answer.*

c) What is the expected number of rainy days X in a 30-day period?

$E[X] = \boxed{}$

d) What is the variance of the number of rainy days X in a 30-day period?

$Var[X] = \boxed{}$

Name: ID#:

Problem 3 The Lionfish Invasion. *(20 points)*

Lionfish are an invasive species that threaten Florida's native marine wildlife. Marine biologists are conducting a survey to determine whether each site in the survey area requires human intervention. Volunteer divers are asked to report the number of lionfish they encounter at each dive site.

If a site does not require intervention ($Y = 0$), a diver may encounter between 0 and 17 lionfish, $0 \leq X \leq 17$, with equal probabilities. If a site requires intervention ($Y = 1$), they may encounter between 11 and 50 lionfish, $11 \leq X \leq 50$, with equal probabilities. Before the divers conduct their survey, marine biologists assume that $P(Y = 1) = 0.4$.

When computing probabilities or expected values below, **enter the numerical value of your answer in the provided box**, and also include equations showing how you calculated that answer.

a) *What is $P(Y = 1 \mid X = 20)$, the probability that the site requires intervention given that 20 lionfish were observed at a site?*

$$P(Y = 1 \mid X = 20) = \boxed{}$$

b) *What is $P(Y = 1 \mid X = 12)$, the probability that the site requires intervention given that 12 lionfish were observed at a site?*

$$P(Y = 1 \mid X = 12) = \boxed{}$$

c) *What is the expected number of lionfish observed at a site?*

$E[X] = \boxed{}$

d) *For which lionfish counts* $0 \leq x \leq 50$ *is the posterior probability* $P(Y = 1 \mid X = x) > 0.5$*? Justify your answer.*

Name: [ ] ID#: [ ]

Problem 4 Truth and Falsehood. *(5 points)*

For each statement below, please **completely fill in the bubble** next to either **True** or **False**. Only select **True** if a statement is *always* true, for any distribution of the referenced random variables. You do *not* need to explain or justify your answers.

a) $E[X^2] \geq (E[X])^2$

◯ **True** ◯ **False**

b) $E[X + Y] = E[X] + E[Y]$

◯ **True** ◯ **False**

c) $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$

◯ **True** ◯ **False**

d) If X and Y are independent, then they are also conditionally independent given $Z = z$.

◯ **True** ◯ **False**

e) Consider a hash table with $N = 2^{64} \approx 10^{19}$ bins. If only $N/1000$ items have been stored in the table, then the probability of a collision (two items being hashed to the same bin) is small.

◯ **True** ◯ **False**

This page is intentionally blank, use as you wish.

Common Discrete Random Variables

Geometric For a geometric variable X with success probability $0 < q \leq 1$,

$$\begin{aligned}
p_X(x) &= (1-q)^{x-1}q, \qquad x \in \{1,2,3,\ldots\}.\\
F_X(x) = P(X \leq x) &= 1-(1-q)^x, \qquad x \in \{1,2,3,\ldots\}.\\
E[X] &= \sum_{x=1}^{\infty} x p_X(x) = \frac{1}{q}.\\
\mathrm{Var}[X] &= E[X^2] - E[X]^2 = \frac{1-q}{q^2}.
\end{aligned}$$

Binomial For a binomial variable X with n trials and success probability $0 \leq q \leq 1$,

$$\begin{aligned}
p_X(x) &= \binom{n}{x} q^x (1-q)^{n-x} = \frac{n!}{(n-x)!x!} q^x (1-q)^{n-x}, \qquad x \in \{0,1,\ldots,n\}.\\
E[X] &= \sum_{x=0}^{n} x p_X(x) = nq.\\
\mathrm{Var}[X] &= E[X^2] - E[X]^2 = nq(1-q).
\end{aligned}$$

Uniform For a random variable X with a discrete uniform distribution between a and b,

$$\begin{aligned}
p_X(x) &= \begin{cases} 1/(b-a+1), & \text{if } x \in \{a, a+1, \ldots, b-1, b\}.\\ 0, & \text{otherwise.} \end{cases}\\
E[X] &= \sum_{x=-\infty}^{\infty} x p_X(x) = \frac{a+b}{2}.\\
\mathrm{Var}[X] &= E[X^2] - E[X]^2 = \frac{(b-a+1)^2 - 1}{12}.
\end{aligned}$$

Summations and Series

$$\begin{aligned}
\sum_{i=0}^{\infty} q^i &= \frac{1}{1-q}, \qquad \text{if } 0 < q < 1.\\
\sum_{i=1}^{n} i &= \frac{n(n+1)}{2}.
\end{aligned}$$