Q1 (25 pts): Theoretical Questions

1. (8 pts) True or False *For full credits, include a justification or example in your answer.*

(a) A continuous distribution's density can take values exceeding 1 at some points.

TRUE, for example, $f(x) = \begin{cases} 2, & \text{if } 0 \leq x \leq 0.5, \\ 0, & \text{if otherwise} \end{cases}$

So, $\int_{-\infty}^{\infty} f(x) = 1$ and $f(x) \geq 0$

(b) For a real-valued random variable X with $F(t) = P(X \leq t)$, it is possible for the cumulative distribution function to exceed 1.

FALSE, by definition the cumulative distribution is mapping with a range of $[0, 1]$

(c) Suppose a pair of real random variables has joint density

$$f(X, Y) = XY; \quad 0 \leq X \leq 1, \quad 0 \leq Y \leq 2.$$

Then X and Y are independent.

$f(X) = \int_0^2 XY \, dY = 2X$

$f(Y) = \int_0^1 XY \, dY = \frac{1}{2}Y$

$f(X) \cdot f(Y) = XY$

so, TRUE

(d) If A and B are conditionally independent given Z, meaning $P(A, B \mid Z) = P(A \mid Z)P(B \mid Z)$, then conditioning further on B does not change A's conditional probability: $P(A \mid B, Z) = P(A \mid Z)$.

$$P(A|B, Z) = \frac{P(B, Z|A)P(A)}{P(B \cap Z)} \text{ by Baye's}$$

$$= \frac{P(B|Z)P(Z|A)P(A)}{P(B \cap Z)} \text{ because } A \text{ and } B \text{ conditionally independent.}$$

$$= \frac{\frac{P(B \cap Z)}{P(Z)}\frac{P(Z \cap A)}{P(A)}P(A)}{P(B \cap Z)} = \frac{P(Z \cap A)}{P(Z)} = P(A|Z)$$

so it must be TRUE.

2. (4 pts) A reliability study tracks the lifetimes (in months) of 60 identical batteries: 5 fail in $[60, 69]$, 25 in $[70, 79]$, 20 in $[80, 89]$, and 10 in $[90, 99]$. Using this sample, estimate the following quantities

(a) What is the estimated probability $P(A)$ that a randomly chosen battery lasts more than 69 months?

$$P(A) = 1 - \frac{5}{60} = 1 - \frac{1}{12} = \frac{11}{12}$$

(b) What is the estimated probability $P(B)$ that a randomly chosen battery lasts more than 79 months?

$$P(B) = \frac{10}{60} + \frac{20}{60} = \frac{1}{2}$$

3. (4 pts) An box contains an equal number of red and blue balls. Draw with replacement repeatedly until the first red ball appears.

(a) What is the sample space for this experiment? What is the probability that the first red appears on the i-th draw?

$\{R, BR, BBR, B..BR, ...\}$

$P(X = i) = P(\text{blue}) \cdot P(\text{red})^{i-1} = (\frac{1}{2})(\frac{1}{2})^{i-1} = (\frac{1}{2})^i$

(b) Let E be the event that the first red appears after an even number of draws. Which outcomes constitute E? What is $P(E)$?

$\{BR, BBBR, BBBBBR, ...\}$

$P(X = 2i - 1) = \frac{1}{2} \cdot (\frac{1}{2})^{2i-1} = (\frac{1}{4})^i$

$$P(E) = \sum (\frac{1}{4})^i = \frac{1}{3}$$

4. (4 pts) Determine the unique matrix A^{-1} satisfying $AA^{-1} = I$ for

$$A = \begin{bmatrix} 5 & 4 & -7 \\ 4 & 3 & -5 \\ 2 & 4 & 1 \end{bmatrix}$$

$det(A) = 11$

so, $A^{-1} = \dfrac{1}{11}\begin{bmatrix} -23 & 32 & -1 \\ 14 & -19 & 3 \\ -10 & 12 & 1 \end{bmatrix}$

5. (5 pts) Independence

A small game is played with three fair, independent "switches." Each switch $i \in \{1, 2, 3\}$ outputs a bit $T_i \in \{0, 1\}$ with $P(T_i = 1) = P(T_i = 0) = \frac{1}{2}$, independently across i (i.e., $T_i \sim \text{Bernoulli}(\frac{1}{2})$). The game then lights up three indicators defined by

$$X = T_1 \oplus T_2, \qquad Y = T_2 \oplus T_3, \qquad Z = T_3 \oplus T_1,$$

where $\oplus$ denotes the XOR (exclusive OR) operator. Show that the random variables X, Y, Z are pairwise independent but not mutually independent.

$T_1 = 1 \oplus T_2 = 1 \Rightarrow X = 0$

$T_1 = 1 \oplus T_2 = 0 \Rightarrow X = 1$

$T_1 = 0 \oplus T_2 = 1 \Rightarrow X = 1$

$T_1 = 0 \oplus T_2 = 0 \Rightarrow X = 0$

so, $P(X = 0) = P(X = 1) = \dfrac{1}{2}$

Similarly, $P(Y = 0) = P(Y = 1) = \dfrac{1}{2}$, and $P(Z = 0) = P(Z = 1) = \dfrac{1}{2}$

so, for $P(X = 0, Y = 0) = \dfrac{2}{8} = \dfrac{1}{4} = \dfrac{1}{2} \cdot \dfrac{1}{2} = P(X = 0) \cdot P(Y = 0)$

because out of the possible 8 cases $X = 0, Y = 0$ occurs when $T_1 = T_2 = T_3 = 0$ or 1

and same for $P(X = 1, Y = 0) = P(X = 0, Y = 1) = P(X = 0, Y = 0) = \dfrac{1}{4}$ (with their respective 2 out of 8 cases)

So we see that X and Y are pairwise independent and a similar case is made for X and Z, Y and Z

They are not mutually independent because, for example,

$P(X = 1, Y = 0, Z = 1) = \dfrac{1}{4} \neq \dfrac{1}{2} \cdot \dfrac{1}{2} \cdot \dfrac{1}{2} = P(X = 1) \cdot P(Y = 0) \cdot P(Z = 0)$

Q3 (25 pts): Yelp Dataset