Fall 2022, MATH 3215-J, Exam 2 (30 pts)

Name:

GT ID:

- Regular exam time is 75 min.
- Open-book/notes. Calculators are allowed. No communication in any form.
- Please clearly indicate what each formula/number is referring to. For the 1-point parts, it suffices to give a short answer.
- Consider skipping a part if you get stuck somewhere.
- If there is a mistake in a problem statement making it unsolvable, skip it and you will be awarded full points for that problem. To ensure fairness, no questions will be answered for any student during the test.

1. People with flu-like symptoms participate in a screening test. Suppose that 90% of people with flu will test positive, and 5% of people who do not have flu will also test positive.

(a) (2 pts) Suppose that among people who participate in the screening test, 60% actually have flu. What is the probability that a random person tests positive in the screening test?

Let E be the event that a person has flu and let F be the event that a person tests positive. Then

$$\mathbb{P}(F) = \mathbb{P}(F \mid E) \cdot \mathbb{P}(E) + \mathbb{P}(F \mid E^c) \cdot \mathbb{P}(E^c) = 0.9 \cdot 0.6 + 0.05 \cdot 0.4 = 0.56.$$

(b) (2 pts) Suppose that among people who participate in the screening test, 60% actually have flu. Given that a person tests positive, what is the probability that the person has flu?

By the Bayes' rule,

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(F \mid E) \cdot \mathbb{P}(E)}{\mathbb{P}(F)} = \frac{0.9 \cdot 0.6}{0.56} = \frac{27}{28}.$$

(c) (2 pts) This part is separate from the previous parts. Suppose that 39% of people test positive in the screening test. What fraction of people who participate in the test actually have flu?

Suppose that p of people participating in the test have flu. Similar to part (a),

$$\mathbb{P}(F) = \mathbb{P}(F \mid E) \cdot \mathbb{P}(E) + \mathbb{P}(F \mid E^c) \cdot \mathbb{P}(E^c) = 0.9 \cdot p + 0.05 \cdot (1 - p) = 0.39,$$

giving $p = 0.4$.

2. Suppose that (X, Y) has joint PDF

$$f(x, y) = c\,x \qquad \text{for } x, y \in (0, 1),\ x + y < 1,$$

and $f(x, y) = 0$ otherwise, where c is a constant.

(a) (2 pts) What is the constant c?
We have $1 = \int_0^1 \int_0^{1-x} c\,x\,dy\,dx = \int_0^1 c\,x(1-x)\,dx = c(1/2 - 1/3) = c/6$, so $c = 6$.

(b) (2 pts) Find the marginal PDF $f_Y(y)$ of Y. (Clearly specify the domain of y where the PDF is nonzero.)
We have $f_Y(y) = \int_0^{1-y} 6x\,dx = 3(1-y)^2$ for $0 < y < 1$ and 0 otherwise.

(c) (2 pts) Find the conditional PDF $f_{Y|X}(y|x)$ of Y given $X = x$ where $0 < x < 1$. (Clearly specify the domain of y where the PDF is nonzero.)
We have $f_X(x) = \int_0^{1-x} 6x\,dy = 6x(1-x)$ for $0 < x < 1$, so $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{1}{1-x}$ if $0 < y < 1 - x$ and 0 otherwise.

(d) (1 pt) Are X and Y independent?
No.

3. Consider random variables X and Y taking values in $\{0,1\}$ such that $\mathbb{P}\{X=0,\, Y=0\} = 0.5$ and $\mathbb{P}\{X=0,\, Y=1\} = \mathbb{P}\{X=1,\, Y=0\} = 0.1$.

(a) (1 pt) Let $f(x,y)$ denote the joint PMF of (X,Y). Find $f(1,1)$.
We have $f(1,1) = \mathbb{P}\{X=1,\, Y=1\} = 1 - 0.5 - 0.1 \cdot 2 = 0.3$.

(b) (2 pts) Compute the marginal mean $\mathbb{E}[X]$ and the marginal variance $\mathrm{Var}(X)$ of X.
Marginally, $\mathbb{P}\{X=1\} = 0.1 + 0.3 = 0.4$, so $\mathbb{E}[X] = 0.4$ and $\mathrm{Var}(X) = 0.4(1-0.4) = 0.24$.

(c) (2 pts) Compute the covariance $\mathrm{Cov}(X,Y)$.
We have $\mathbb{E}[XY] = 0.3$, so $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y] = 0.3 - 0.4^2 = 0.14$.

(d) (1 pt) Is the correlation between X and Y positive or negative?
Positive.

(e) (1 pt) Are X and Y independent?
No.

4. The annual rainfall in inches in a region has distribution $\mathcal{N}(30, 10)$, and the rainfalls in different years are independent. (Although the amount of rainfall cannot be negative, you may ignore this minor inaccuracy here because a $\mathcal{N}(30, 10)$ random variable is negative with an extremely small probability.)

(a) (2 pts) What is the probability that the total rainfall is at most 320 inches in a given decade (i.e., 10 years)? Express the answer in the form $\Phi(\dots)$ where Φ is the CDF of $\mathcal{N}(0, 1)$.

The total rainfall in a decade has distribution $X \sim \mathcal{N}(300, 100)$, so the probability is

$$\mathbb{P}\{X \leq 320\} = \mathbb{P}\left\{\frac{X - 300}{\sqrt{100}} \leq \frac{320 - 300}{\sqrt{100}}\right\} = \mathbb{P}\{Z \leq 2\} = \Phi(2),$$

where $Z \sim \mathcal{N}(0, 1)$.

(b) (2 pts) What is the probability that the rainfall will exceed 320 inches in at most 2 of the next 3 decades? Express the answer using the CDF Φ of $\mathcal{N}(0, 1)$, but the final answer needs to be simplified as much as possible (no binomial coefficient, etc).

Suppose Y of the next 3 decades have rainfall exceeding 320 inches. Then $Y \sim \mathrm{Bin}(3, 1 - \Phi(2))$. As a result, the probability is

$$\mathbb{P}\{Y \leq 2\} = \binom{3}{0}\Phi(2)^3 + \binom{3}{1}(1-\Phi(2))\Phi(2)^2 + \binom{3}{2}(1-\Phi(2))^2\Phi(2) = \Phi(2)^3 - 3\Phi(2)^2 + 3\Phi(2).$$

5. Let $X_1, X_2, \ldots, X_{25}$ be i.i.d. random variables from a distribution with mean 10.

(a) (2 pts) Suppose that the random variables take nonnegative values. Use Markov's inequality to give an upper bound on

$$\mathbb{P}\{X_1 + X_2 + \cdots + X_{25} > 1000\}.$$

Since $\mathbb{E}[X_1 + \cdots + X_{25}] = 25 \cdot 10 = 250$, Markov's inequality implies that

$$\mathbb{P}\{X_1 + \cdots + X_{10} > 1000\} \leq \frac{250}{1000} = \frac{1}{4}.$$

(b) (2 pts) Suppose that each X_i has variance 4. Use Chebyshev's inequality to give a lower bound on

$$\mathbb{P}\{-10 \leq X_{10} - X_{20} \leq 10\}.$$

Since $\mathbb{E}[X_{10} - X_{20}] = 0$ and $\mathrm{Var}(X_{10} - X_{20}) = 8$, Chebyshev's inequality implies that

$$\mathbb{P}\{|X_{10} - X_{20}| > 10\} \leq \frac{8}{100} = 0.08, \qquad \mathbb{P}\{-10 \leq X_{10} - X_{20} \leq 10\} \geq 0.92.$$

(c) (2 pts) Suppose that $X_i \sim \mathcal{N}(10, 4)$. Use the CDF Φ of $\mathcal{N}(0,1)$ to express the probability

$$\mathbb{P}\{(X_1 + X_2 + \cdots + X_{10}) - (X_{11} + X_{12} + \cdots + X_{25}) < -100\}.$$

Since

$$\mathbb{E}[(X_1 + \cdots + X_{10}) - (X_{11} + \cdots + X_{25})] = 10 \cdot (10 - 15) = -50$$

and

$$\mathrm{Var}((X_1 + \cdots + X_{10}) - (X_{11} + \cdots + X_{25})) = 4 \cdot 25 = 100,$$

we have

$$\begin{aligned}
&\mathbb{P}\{(X_1 + \cdots + X_{10}) - (X_{11} + \cdots + X_{25}) < -100\} \\
&= \mathbb{P}\left\{\frac{(X_1 + \cdots + X_{10}) - (X_{11} + \cdots + X_{25}) + 50}{\sqrt{100}} < \frac{-100 + 50}{\sqrt{100}}\right\} \\
&= \mathbb{P}\{Z < -5\} = \Phi(-5),
\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$.