MATH 201 (SUMMER '23, SESH A2)

FINAL EXAM
SOLUTIONS

(A) **2. (15(A) points)**

Let X be a continuous random variable with cumulative distribution function (c.d.f.)

$$F_X(x) = \begin{cases} c(2 - e^{-x} - e^{-2x}) & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

where c is a constant.

(a) Find c.

SINCE F_X IS A c.d.f. IT MUST SATISFY

$$F_X(\infty) = \lim_{x \to \infty} F_X(x) = 1$$

CLEARLY, $\lim_{x \to \infty} F_X(x) = \lim_{x \to \infty} c(2 - e^{-x} - e^{-2x}) = 2c \Rightarrow c = \frac{1}{2}$

(b) Find the probability density function (p.d.f.), f_X of X.

SINCE X IS CONT.,

$$f_X(x) = F_X'(x) = \begin{cases} 0 & \text{IF } x < 0 \\ \frac{d}{dx}\left[\frac{1}{2}\left(2 - e^{-x} - e^{-2x}\right)\right] = \frac{e^{-x}}{2} + e^{-2x} & \text{IF } x \geq 0 \end{cases}$$

(c) Find $\mathbb{E}[X]$.

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx$$

$$= \int_{0}^{\infty} x \left(\frac{e^{-x}}{2} + e^{-2x} \right) dx$$

(INT. BY PARTS)

$$= x \left[-\frac{e^{-x}}{2} - \frac{e^{-2x}}{2} \right]_{x=0}^{x=\infty} - \int_{0}^{\infty} 1 \cdot \left[-\frac{e^{-x}}{2} - \frac{e^{-2x}}{2} \right] dx$$

$$= \int_0^\infty \left(\frac{e^{-x}}{2} + \frac{e^{-2x}}{2} \right) dx$$

$$= \left. \frac{-e^{-x}}{2} - \frac{e^{-2x}}{4} \right]_{x=0}^{x=\infty}$$

$$= \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

(A) **3. (15(A) points)** Scientists are trying to do an experiment. They know the outcome of the experiment X is random variable with distribution Ber(1/3). However, due to inaccuracies in their measuring equipment, there is a Gaussian noise parameter $G \sim \mathcal{N}(0,1)$, and what actually gets measured by the equipment is $M = X + G$. Assuming that X and G are independent, what is the probability that $X = 1$ if the measurement M is at most 4?

You may leave your answer in terms of $\Phi(x)$, without evaluating Φ in **this question only**.

WE WANT:

$$P(X=1 \mid M \leq 4)$$

BY BAYES' FORMULA, SINCE $X \in \{0,1\}$,

$$P(X=1 \mid M \leq 4) = \frac{P(X=1, M \leq 4)}{P(M \leq 4)}$$

$$= \frac{P(X=1, M \leq 4)}{P(X=1, M \leq 4) + P(X=0, M \leq 4)}$$

NOW,
$$P(X=a, M\leq 4) = P(X=a, X+G\leq 4)$$
$$= P(X=a, G\leq 4-a)$$
$$= P(X=a)\,P(G\leq 4-a) \quad (\text{IMP. OF } X, G)$$

$$\therefore\ P(X=0, M\leq 4) = P(X=0)\,P(G\leq 4) = \frac{2}{3}\ \Phi(4)$$

$$P(X=1, M\leq 4) = P(X=1)\ P(G\leq 3) = \frac{1}{3}\ \Phi(3)$$

$$\left[\ \because\ X\sim Ben(1/3),\ G\sim N(0,1)\right]$$

$$\therefore P(X=1 \mid M \le 4) = \frac{\frac{1}{3}\Phi(3)}{\frac{1}{3}\Phi(3) + \frac{2}{3}\Phi(4)} = \frac{\Phi(3)}{\Phi(3) + 2\Phi(4)}$$

(A) **4.** **(10(A) points)** In the city of Gotham, anyone who likes drinking tea doesn't like drinking anything else. If everyone likes at least one of the beverages tea, coffee, or soda, 20% of people like tea, 75% people like soda, and 60% of people like coffee, then how many people like both coffee and soda?

SEE SOLNS OF SAMPLE FINAL FOR ONE METHOD.

ALTERNATE : LET $T/C/S$ BE EVENTS THAT A RANDOMLY SAMPLED PERSON LIKES TEA/COFFEE/SODA

PINK TEXT IMPLIES THAT $C \cup S = T^c$

$$\Rightarrow \quad P(C \cup S) = 1 - P(T) = 1 - \frac{20}{100} = 0.8$$

OTOH, $P(C) = \frac{60}{100} = 0.6$, $P(S) = \frac{75}{100} = 0.75$

$\therefore$ BY INCL/EXCL.,

$$P(C \cup S) = P(C) + P(S) - P(C \cap S)$$

$$\Rightarrow P(C \cap S) = P(C) + P(S) - P(C \cup S)$$

$$= 0.6 + 0.75 - 0.8 = 0.55$$

$\therefore$ 55% LIKE BOTH COFFEE & SODA.

(A) **5. (10(A) points)**

A national vote was held about pineapple on a pizza. The result says 40% of the population love pineapple on a pizza, while the others hate it.

(a) Triomino's sells pizzas with pineapple on them at \$15 per pizza, and pizzas without pineapple on them at \$14. Let X_n be the money Alfredo's makes from selling n pizzas. Express X_n in terms of a distribution you know.

LET Y_n = # OF PINEAPPLE PIZZAS SOLD

$\Rightarrow Y_n \sim Bin(n, 0.4)$

$\therefore X_n = 15Y_n + 14(n - Y_n) = 14n + Y_n$

(b) Find

$$\lim_{n\to\infty} \mathbb{P}(X_n > 14.5n).$$

$$P(X_n > 14.5n) = P(14n + Y_n > 14.5n) = P(Y_n > 0.5n)$$

$$= P\left(\frac{Y_n}{n} - 0.4 > 0.1\right)$$

$$\leq 1 - \boxed{P\left(\left|\frac{Y_n}{n} - 0.4\right| \leq 0.1\right)} \to 1 \text{ AS } n \to \infty$$

(BY LAW OF LARGE NUMBERS

$$\therefore \lim_{n \to \infty} P(X_n > 14.5n) \leq 1 - 1 = 0$$

$$\Rightarrow \lim_{n \to \infty} P(X_n > 14.5n) = 0$$

(B) **6. (30(B) points)**

Suppose (X, Y) are uniformly distributed on the triangular region

$$D = \{(x, y) : x + y \leq 1, x, y \geq 0\}$$

SEE SOLN. OF SAMPLE FINAL

ANSWERS FOR (C):

$$\mathbb{E}[X] = \mathbb{E}[Y] = 1/3$$

$$\mathrm{Var}[X] = \mathrm{Var}[Y] = 1/18$$

$$\mathrm{Cov}[X, Y] = -1/36$$

$$\mathrm{Corr}[X, Y] = -1/2$$

(B) **7. (30(B) points)**

Suppose X and Y are independent random variables with the following moment generating functions (m.g.f.),

$$M_X(t) = \frac{1}{4}e^{-t} + \frac{3}{4}e^{t},$$

$$M_Y(t) = \frac{2}{3} + \frac{1}{3}e^{2t}.$$

(a) Find the joint p.m.f., $p_{X,Y}(x, y)$ and express it as a table.

X \ Y		0	2
	p_X \ p_Y	(2/3)	(1/3)
-1	(1/4)	1/6	1/12
1	(3/4)	1/2	1/4

(b) Let $Z = X + Y$. What is $M_Z(t)$?

$$M_Z(t) = M_X(t)\, M_Y(t)$$

$$= \left(\frac{1}{4} e^{-t} + \frac{3}{4} e^{t}\right)\left(\frac{2}{3} + \frac{1}{3} e^{2t}\right)$$

$$= \frac{1}{6} e^{-t} + \frac{7}{12} e^{t} + \frac{1}{4} e^{3t}$$

(c) Compute the p.m.f. p_Z.

BY INSPECTION OF $M_Z(t)$, $Z \in \{-1, 1, 3\}$

$$p_Z(-1) = \frac{1}{6}$$

$$p_Z(1) = \frac{7}{12}$$

$$p_Z(3) = \frac{1}{4}$$

(d) Compute $\mathbb{E}[Z^3] - E[Z^2]$.

$$= M_Z^{(3)}(0) - M_Z^{(2)}(0)$$

$$= \frac{43}{6} - 3 = \frac{25}{6}$$

(B) **8. (15(B) points)**

Recall that NegBin(k,p) for $k \in \mathbb{N}$ and $0 < p < 1$ is the negative binomial distribution and it is defined as the number of independent Ber(p) trials before one sees k successes.

The p.m.f. of $X \sim \text{NegBin}(k,p)$ is given by

$$p_X(n) = P(X = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}.$$

In the following parts, it may be useful to recall that if $Y \sim \text{Geom}(p)$, then

$$E[Y] = \frac{1}{p} \qquad \text{Var}[Y] = \frac{1-p}{p^2}.$$

(a) Find a formula for $\mathbb{E}[X]$ and Var$[X]$. (**Hint:** do not compute this directly! Instead, use the relationship between NegBin and Geom.)

SEE SOLN. OF SAMPLE FINAL.

(b) If $X_1 \sim \text{NegBin}(k_1, p)$ and $X_2 \sim \text{NegBin}(k_2, p)$ are independent, then what is the distribution of $X_1 + X_2$? (**Hint:** there is a simple way to do this without using convolutions!)

$X_1 + X_2$ = # OF INDEPENDENT $Ber(p)$ TRIALS BEFORE $k_1 + k_2$ SUCCESSES WERE OBSERVED

$\sim \text{NegBin}(k_1 + k_2, p)$

(B) **9. (20(B) points)**

On average Batman and Robin have to wait 36 hours between the occurrences of two crimes important enough for Commissioner Gordon to light up the Bat-Signal. The Bat-Signal was just lit up, and suppose B is the number of hours from now until the next time the Bat-Signal is lit up.

(a) If you know nothing else about the distribution of B, what upper bound can you provide for the probability that $B > 90$? Clearly state any inequality you use, and explain why the hypotheses apply.

CLEARLY, $B \geq 0$ SO MARKOV'S INEQUALITY APPLIES

$$\Rightarrow P(B > 90) \leq \frac{E(B)}{90} = \frac{36}{90} = \frac{2}{5}$$

(b) You can assume B is a continuous random variable. Further, B is memoryless – if no Bat-signal worthy crime has occurred in the t hours since the last one, then probability it will take at least s more hours is the same as the probability that he would have taken s hours in the first place. That is,

$$P(B > s + t | B > t) = P(B > s).$$

Can you identify the distribution of B?

CONT. + MEMORYLESS $\Rightarrow$ EXPONENTIAL.

$$\mathbb{E}[\text{Exp}(\lambda)] = 1/\lambda = 36 \Rightarrow \lambda = 1/36$$

$$\therefore B \sim \text{Exp}\left(\frac{1}{36}\right)$$

(c) Compute $P(B > 90)$ exactly.

$$= \int_{90}^{\infty} f_B(t)\, dt$$

$$= \int_{90}^{\infty} \lambda e^{-\lambda t}\, dt = e^{-90\lambda} = e^{-90/36}$$

(B) **10.** **(20(B) points)**

Alexander rolls a standard die repeatedly. Let A_j be the indicator random variable for the event that the jth roll is odd, and B_j be the indicator random variable for the event that the jth roll is 5.

(a) Find $\mathbb{E}[A_j B_k]$. (**Hint:** consider the cases $j = k$ and $j \neq k$ separately)

$$\mathbb{E}[A_j B_k] = P[\text{jth ROLL IS ODD AND kth ROLL IS 5}]$$

IF $j = k$,

$$E[A_j B_j] = P[\text{jth ROLL IS 5}] = \frac{1}{6}$$

IF $j \neq k$,

$$\mathbb{E}[A_j B_k] = P(\text{jth ROLL IS ODD}) \cdot P(\text{kth ROLL IS 5})$$

$$= \frac{3}{6} \cdot \frac{1}{6} = \frac{1}{12}$$

$$\therefore \mathbb{E}[A_j B_k] = \begin{cases} 1/12, & j \neq k \\ 1/6, & j = k. \end{cases}$$

(b) Let X be the number of odd numbers and Y be the number of 5s that show up in n rolls of the die. Find $\mathbb{E}[XY]$.

CLEARLY, $X = \sum_{j=1}^{n} A_j$

$$Y = \sum_{k=1}^{n} B_k$$

$$\therefore \mathbb{E}[XY] = \mathbb{E}\left[\sum_{j=1}^{n} \sum_{k=1}^{n} A_j B_k\right] = \sum_{j,k=1}^{n} \mathbb{E}[A_j B_k]$$

LINEARITY OF EXPECTATION.

$$\therefore E[XY] = \frac{n}{6} + \frac{n(n-1)}{12}$$

$j = k$ $\qquad$ $j \neq k$

$$= \frac{n(n+1)}{12}$$

(c) Compute $\mathrm{Cov}[X, Y]$.

CLEARLY,

$$\mathbb{E}[X] = \sum_{j=1}^{n} \mathbb{E}[A_j] = \frac{n}{2}$$

LoE

$$\mathbb{E}[Y] = \sum_{k=1}^{n} \mathbb{E}[B_k] = \frac{n}{6}$$

$$\therefore \mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{n(n+1)}{12} - \left(\frac{n}{2}\right)\left(\frac{n}{6}\right) = \frac{n}{12}.$$

(B) **11. (15(B) points)** No Hobbit in Middle-Earth is taller than 5 feet. You decide to go around randomly asking Hobbits their height to get a good estimate for their average height. Using the law of large numbers, estimate how many Hobbits you need to ask before there is a 99% chance that the average height you sampled differs from the actual average by at most 1 inch? [**Note:** 12 inches is 1 feet.]

Hint: Use Chebyshev's inequality.

SEE SOLN. OF SAMPLE FINAL.