

MTH 201: Probability and Statistics

Mid Semester Exam (Section B)

22/02/2025

Sanjit K. Kaul

No books, notes, or devices are allowed. Just a pen/ pencil and eraser. Institute rules will apply with regards to cheating. Show your steps. Don't prioritize simplifying products and sums, if you are short on time. You have about 120 minutes.

Question 1. 4 marks The previous year $10^{-4}\%$ of people living in a city are hit by lightning. Assume that whether a person is hit or not by lightning is independent of how others are impacted by it. You go to the city to meet its people. The first 5 people you meet have all been hit by lightning in the last year. You meet 5 more people. Calculate the probability that all these 5 people have also been hit by lightning in the previous year.

Question 2. 5 marks Let $\Phi(x)$ be the CDF of the standard normal $N(0, 1)$. Write the probability that a Gaussian random variable with mean 10 and variance 100 takes a value greater than 20 in terms of $\Phi(x)$.

Question 3. 20 marks An AND gate takes two inputs, each of which is chosen uniformly and randomly from the set of values $\{0, 1\}$. We have k AND gates, each of which takes two inputs as described above. Also, inputs taken by AND gates are independent of each other.

Every AND gate outputs the logical AND of its two inputs. The logical AND operation outputs 1 when both inputs to it are 1 and outputs 0 otherwise. Derive the distribution (PMF) of the sum of the outputs of the k AND gates. Is the distribution of a well known family? Which one? Derive the expected value of the sum of the outputs.

Further assume that the number of logical AND gates is chosen from a Poisson distribution with parameter $\alpha = 10$. Derive the distribution and the expected value of the sum of outputs of the AND gates.

Question 4. 36 marks Cars arrive in sequence (one after the other) at a toll plaza that has two toll booths numbered 1 and 2. Assume that the first car to arrive at the plaza chooses either booth 1 or 2 with equal probability. For cars $i = 2, 3, \dots$, car i chooses booth 1 with probability 0.9 if car $i - 1$ chose booth 1 and car i chooses booth 2 with probability 0.9 if car $i - 1$ chose booth 2. To exemplify, if car 1 chooses booth 1, car 2 chooses booth 1 with probability 0.9 and booth 2 otherwise. Similarly, if car 1 chooses booth 2, car 2 chooses booth 2 with probability 0.9 and booth 1 otherwise. Answer the following questions.

- (a) What is the probability that the first car chooses booth 2?
- (b) Derive the distribution (PMF) of the total number of cars that arrive at the toll plaza up to and including the car that is the **first** car to choose booth 2.
- (c) Derive the expected value of the total number of cars that arrive at the toll plaza up to and including the car that is the **first** car to choose booth 2.
- (d) Derive the expected value of the total number of cars that arrive at the toll plaza up to and including the car that is the **second** car to choose booth 2. [Hint: Draw a tree diagram and spot well known random variables. Then try and calculate the expected value directly without first calculating the PMF. Part (c) should be useful.]

Question 5. 35 marks A student enters a lecture and awaits the in-class exercise. At the beginning of a lecture, the instructor sets a timer that expires (times out) after a random amount of time that is distributed as an exponential random variable with mean $1/\lambda = 60$ minutes. The in-class exercise is conducted if the timer expires within the

duration of the lecture, which is 120 minutes long. If the timer expires within the lecture, the in-class exercise is conducted when the timer expires. Answer the following questions.

- 1) Calculate the probability that an in-class exercise is not conducted during a lecture.
- 2) An in-class exercise is conducted during a lecture. Derive the distribution (CDF or PDF) of the time in the lecture at which the in-class exercise is conducted.
- 3) Suppose a student leaves the lecture as soon as the in-class exercise is conducted. In case an in-class exercise is not conducted, the student attends the entire 120 minute lecture. Derive the CDF of the time the student spends in the lecture. Illustrate the CDF (an approximate sketch is enough).

I. USEFUL FACTS

An exponential RV X with expected value (mean) $1/\lambda$ has a PDF $f_X(x) = \lambda e^{-\lambda x}, x \geq 0$.

A Poisson random variable K with parameter α has PMF $P_K(k) = \alpha^k e^{-\alpha} / k!$. Its mean is α .

Question 1. 4 marks The previous year $10^{-4}\%$ of people living in a city are hit by lightning. Assume that whether a person is hit or not by lightning is independent of how others are impacted by it. You go to the city to meet its people. The first 5 people you meet have all been hit by lightning in the last year. You meet 5 more people. Calculate the probability that all these 5 people have also been hit by lightning in the previous year.

$$P[\text{A randomly chosen person was hit by lightning}] = \frac{10^{-4}}{10^2} = 10^{-6}$$

→ Up to $(1/4)$, i.e.

final answer is wrong.

Since being struck by lightning is said to be independent,

$$P[\text{The five more people are struck by lightning} \mid \text{First 5 have been struck}]$$

Up to $(1/4)$, i.e.
final answer is wrong.

$$= P[\text{The five more people are struck by lightning}] = (10^{-6})^5 \\ = 10^{-30}$$

Correct answer with some evidence of understanding
 $(4/4)$
(stating independence, for example, or connect intuition)

Question 2. 5 marks Let $\Phi(x)$ be the CDF of the standard normal $N(0, 1)$. Write the probability that a Gaussian random variable with mean 10 and variance 100 takes a value greater than 20 in terms of $\Phi(x)$.

Suppose X is $\text{Gaussian}(10, \sqrt{100}) = \text{Gaussian}(10, 10)$

$$P[X > 20]$$

$$= P[X - 10 > 20 - 10]$$

$$= P\left[\frac{X - 10}{10} > \frac{20 - 10}{10}\right]$$

$$= 1 - P\left[\frac{X - 10}{10} \leq 1\right]$$

$$= 1 - \Phi(1), \text{ since}$$

$\frac{X - 10}{10}$ is $\text{Gaussian}(0, 1)$,
which is the
standard normal.

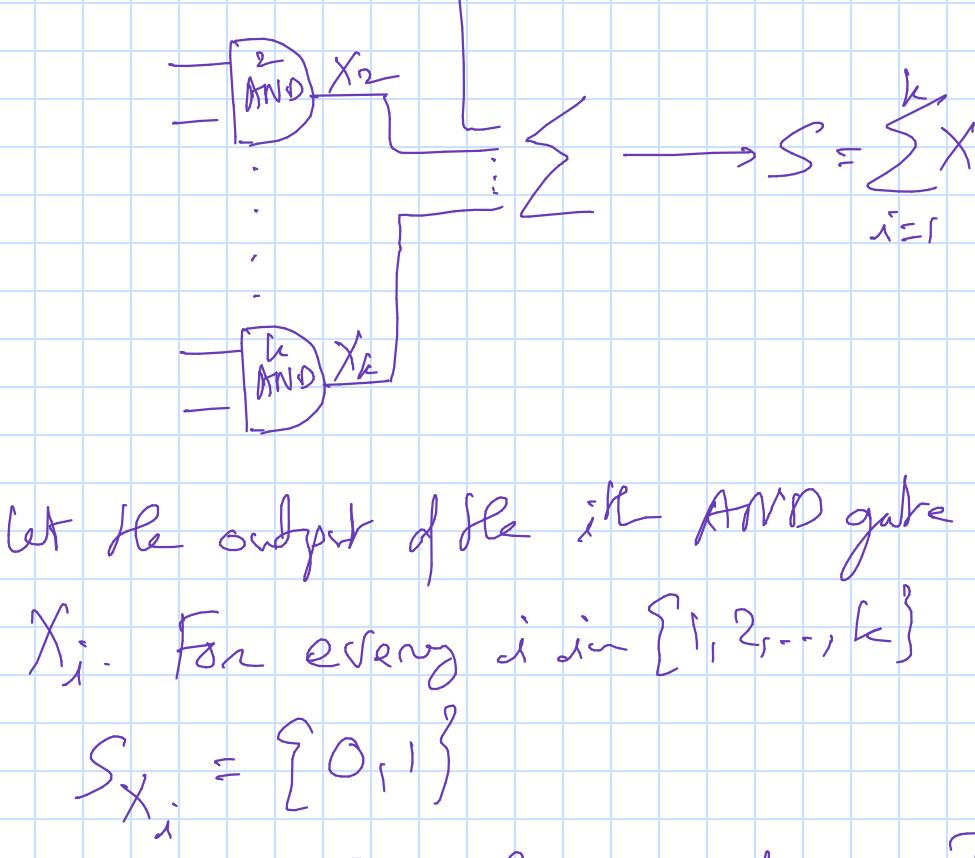
Connect shifting &
scaling of the given
Gaussian \sim — Up to $\frac{2-\cancel{5}}{\cancel{5}}$

Find correct
answer — $\frac{2-\cancel{5}}{\cancel{5}}$

Question 3. 20 marks An AND gate takes two inputs, each of which is chosen uniformly and randomly from the set of values {0, 1}. We have k AND gates, each of which takes two inputs as described above. Also, inputs taken by AND gates are independent of each other.

Every AND gate outputs the logical AND of its two inputs. The logical AND operation outputs 1 when both inputs to it are 1 and outputs 0 otherwise. Derive the distribution (PMF) of the sum of the outputs of the k AND gates. Is the distribution of a well known family? Which one? Derive the expected value of the sum of the outputs.

Further assume that the number of logical AND gates is chosen from a Poisson distribution with parameter $\alpha = 10$. Derive the distribution and the expected value of the sum of outputs of the AND gates.



Let the output of the i^{th} AND gate be X_i .

X_i : For every $i \in \{1, 2, \dots, k\}$

$$S_{X_i} = \{0, 1\}$$

$$P[X_i = 1] = P[\text{Both inputs are } 1] = (0.5)^2 = 0.25.$$

$$P[X_i = 0] = 1 - 0.25 = 0.75.$$

} Stating the PMF and for stating the family of r.v.s X_i . Up to 3/10

Every X_i is a Bernoulli(0.25)

g.v.

Let S be the sum of the outputs of the k AND gates.

$$S = \sum_{i=1}^k X_i$$

S is the sum of k independent

Bern r.v.s. Therefore S is Binomial($k, 0.25$).

} Recognizing that the sum is Binomial. Up to 3/10

$$E[S] = (0.25)k. \quad \left(\begin{array}{l} \text{In case you don't recall} \\ \text{this fact about Binomial r.v.s,} \\ \text{observe that} \end{array} \right)$$

$$\begin{aligned} E[S] &= E[X_1 + X_2 + \dots + X_k] \\ &= E[X_1] + \dots + E[X_k] \\ &= (0.25) + \dots + (0.25) \\ &= (0.25)k. \end{aligned}$$

} Calculating the mean. Up to 3/10

In the above the number of AND gates is a given number k .

To solve the next part of the question, the number of AND gates is a Poisson(10) r.v.

Let K be the no. of AND gates, where

K is Poisson(10).

The range space of K is $\{0, 1, 2, \dots\}$

Consider the PMF of the sum S .

$$P[S=s] = \sum_{k=0}^{\infty} P[S=s | K=k] P[K=k]$$

$$= P[S=s | K=0] P[K=0]$$

$$+ \sum_{k=1}^{\infty} P[S=s | K=k] P[K=k].$$

} Writing the unconditional PMF in terms of the conditional PMF & the Poisson PMF.

Consider $P[S=s | K=0]$.

When $k=0$, we are summing 0 AND gates.

$$P[S=0 | K=0] = 1.$$

} Up to 1/5

For $k \geq 1$,

$$P[S=s | K=k]$$

= $P[s \text{ ones in } k \text{ independent Bernoulli}(0.25) \text{ trials}]$

$$= k C_s (0.25)^s (0.75)^{k-s},$$

$s = 0, 1, 2, \dots, k$.

} Up to 1/5

$$P[S=s] = P[S=s | K=0] P[K=0]$$

$$+ \sum_{k=1}^{\infty} P[S=s | K=k] P[K=k]$$

$$= 0 + \sum_{k=1}^{\infty} (0.25)^k P[K=k]$$

$$= (0.25) \sum_{k=1}^{\infty} k P[K=k]$$

$$= (0.25) E[K] = (0.25)\alpha$$

} Up to 1/5

$$= \underline{0.25}.$$

Question 4. 36 marks Cars arrive in sequence (one after the other) at a toll plaza that has two toll booths numbered 1 and 2. Assume that the first car to arrive at the plaza chooses either booth 1 or 2 with equal probability. For cars $i = 2, 3, \dots$, car i chooses booth 1 with probability 0.9 if car $i-1$ chose booth 1 and car i chooses booth 2 with probability 0.9 if car $i-1$ chose booth 2. To exemplify, if car 1 chooses booth 1, car 2 chooses booth 1 with probability 0.9 and booth 2 otherwise. Similarly, if car 1 chooses booth 2, car 2 chooses booth 2 with probability 0.9 and booth 1 otherwise. Answer the following questions.

(a) What is the probability that the first car chooses booth 2? **6/36**

(b) Derive the distribution (PMF) of the total number of cars that arrive at the toll plaza up to and including the car that is the **first** car to choose booth 2. **10/36**

(c) Derive the expected value of the total number of cars that arrive at the toll plaza up to and including the car that is the **first** car to choose booth 2. **10/36**

(d) Derive the expected value of the total number of cars that arrive at the toll plaza up to and including the car that is the **second** car to choose booth 2. [Hint: Draw a tree diagram and spot well known random variables.]

Then try and calculate the expected value directly without first calculating the PMF. Part (c) should be useful.]

$$(a) P[\text{first car chooses booth 2}] = 0.5.$$

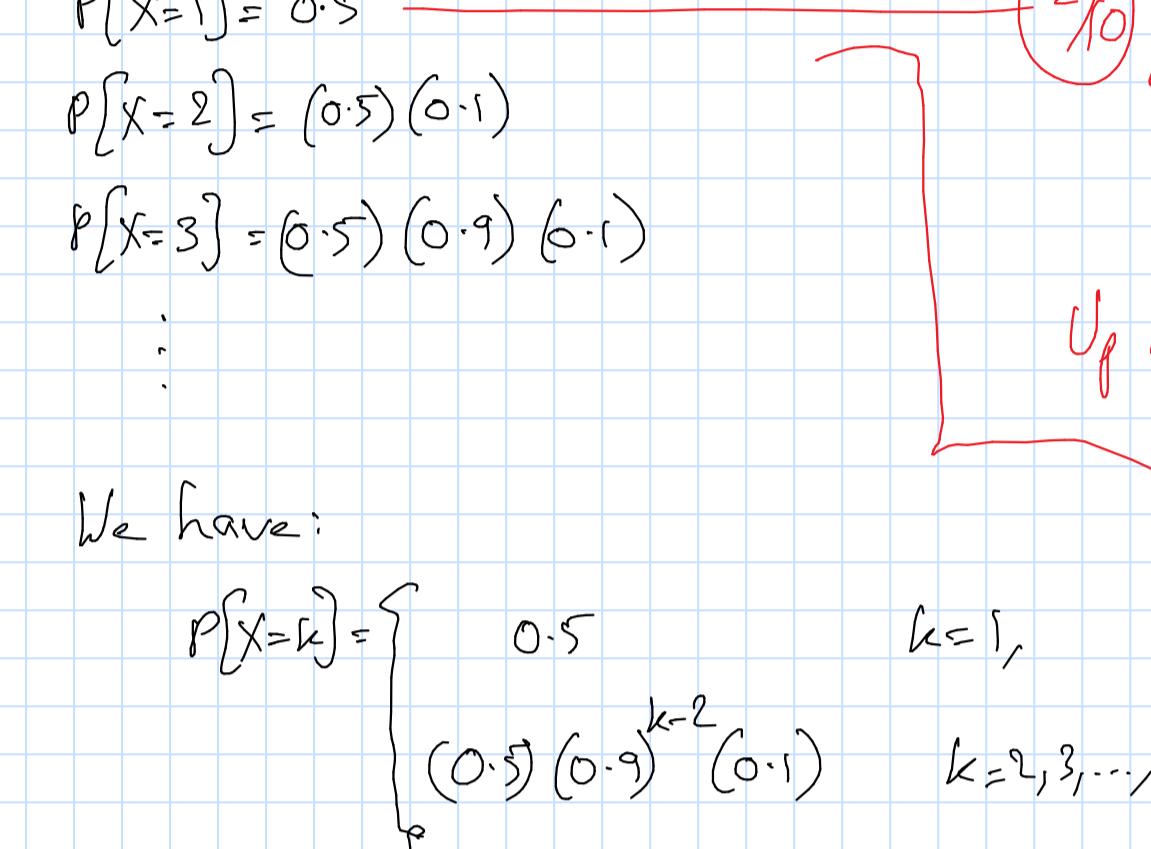
This is given in the question.

6/6 for the
correct answer.

$$(b) \text{Let } (i, 1) \text{ denote car } i \text{ choosing booth 1.}$$

$$\text{Let } (i, 2) \text{ denote car } i \text{ choosing booth 2.}$$

Let X be the no. of cars that arrive at the toll plaza up to and including the first car to choose booth 2.



Up to 9 for a decent attempt if PMF is wrong
After all which includes free diagrams etc.

$$P[X=1] = 0.5$$

$$P[X=2] = (0.5)(0.1)$$

$$P[X=3] = (0.5)(0.9)(0.1)$$

:

We have:

$$P[X=k] = \begin{cases} 0.5 & k=1, \\ (0.5)(0.9)^{k-2}(0.1) & k=2, 3, \dots, \\ 0 & \text{otherwise.} \end{cases}$$

Up to 8/10 for the next.

$$(c) E[X]$$

$$= (0.5)1 + (0.5)(0.1)(2)$$

$$+ (0.5)(0.9)(0.1)(3)$$

$$+ (0.5)(0.9)^2(0.1)(4) + \dots$$

$$= 0.5 + (0.5)(0.1) \left[2 + (0.9)3 + (0.9)^24 + \dots \right]$$

$$= 0.5 + \frac{(0.5)}{0.9} \left[(0.9)(0.1)2 + (0.9)^2(0.1)3 + (0.9)^3(0.1)4 + \dots \right]$$

$$= 0.5 + \frac{0.5}{0.9} \frac{(0.9)(1.1)}{(0.9)(0.1)}$$

$$= 0.5 + 0.5(1.1) = 1.1(0.5) = 6.1$$

8/10 for correctly expressing $E[X]$.

2/10 for final answer

(d) Let Y count the no. of cars that arrive at the toll plaza after the first car that chose Toll 2, up to and including the first car amongst the arrivals to choose Toll 2.

Note that $X+Y$ is the total no. of cars that arrive at the toll plaza up to and including the second car to choose Toll 2.

Switely breaking down the problem up for **31/10**

Note that Y counts cars after the first car to choose Toll 2.

The Second car to choose Toll 2 arrived.

First car to choose Toll 2 arrived.

Toll 1 is chosen.

Toll 2 is chosen.

Toll 1 is chosen.

Question 5. 35 marks A student enters a lecture and awaits the in-class exercise. At the beginning of a lecture, the instructor sets a timer that expires (times out) after a random amount of time that is distributed as an exponential random variable with mean $1/\lambda = 60$ minutes. The in-class exercise is conducted if the timer expires within the duration of the lecture, which is 120 minutes long. If the timer expires within the lecture, the in-class exercise is conducted when the timer expires. Answer the following questions.

- 1) Calculate the probability that an in-class exercise is not conducted during a lecture. 9/35
- 2) An in-class exercise is conducted during a lecture. Derive the distribution (CDF or PDF) of the time in the lecture at which the in-class exercise is conducted. 15/35
- 3) Suppose a student leaves the lecture as soon as the in-class exercise is conducted. In case an in-class exercise is not conducted, the student attends the entire 120 minute lecture. Derive the CDF of the time the student spends in the lecture. Illustrate the CDF (an approximate sketch is enough). 15/35

i) $P[\text{In-class exercise is not conducted during a lecture}]$

$$= P[\text{Timer expires after 120 minutes}].$$

Let X be the time the timer expires.

We require

$$P[X > 120] = e^{-\left(\frac{1}{60}\right)(120)} = e^{-2},$$

Since X is exponential ($\frac{1}{60}$).

2) We have been told that an in-class exercise was conducted. That is $\{X \leq 120\}$.

The distribution of X given $\{X \leq 120\}$ is

{ Merely recognizing that we require a conditional distribution } (3/15)

Up to

↓
Construct CDF/CCDF/PDF

With the needed math (as described) (15/35)

$$P[X > x | X \leq 120]$$

$$= \begin{cases} \frac{P[X > x, X \leq 120]}{P[X \leq 120]} & 0 \leq x \leq 120 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{P[x < X \leq 120]}{P[X \leq 120]} & 0 \leq x \leq 120 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{P[X \leq 120] - P[X \leq x]}{P[X \leq 120]} & 0 \leq x \leq 120 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{(1 - e^{-2}) - (1 - e^{-x/60})}{1 - e^{-2}} & 0 \leq x \leq 120 \\ 0 & \text{otherwise} \end{cases}$$

{ It is okay to state the CCDF. Don't have to calculate the CDF or the PDF. (15/35)

The above is the conditional complementary CDF. The conditional CDF is

$$P[X \leq x | X \leq 120] = \begin{cases} \frac{1 - e^{-x/60}}{1 - e^{-2}} & 0 \leq x \leq 120, \\ 0 & \text{otherwise.} \end{cases}$$

If you prefer the PDF route, we know

Let

$$f_{X|X \leq 120}(x) = \begin{cases} \frac{f_X(x)}{P[X \leq 120]} & 0 \leq x \leq 120, \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} \frac{\left(\frac{1}{60}\right)e^{-x/60}}{1 - e^{-2}} & 0 \leq x \leq 120, \\ 0 & \text{otherwise.} \end{cases}$$

3) CDF of the time a student spends in the lecture. Let Y be the time.

The student spends 120 minutes

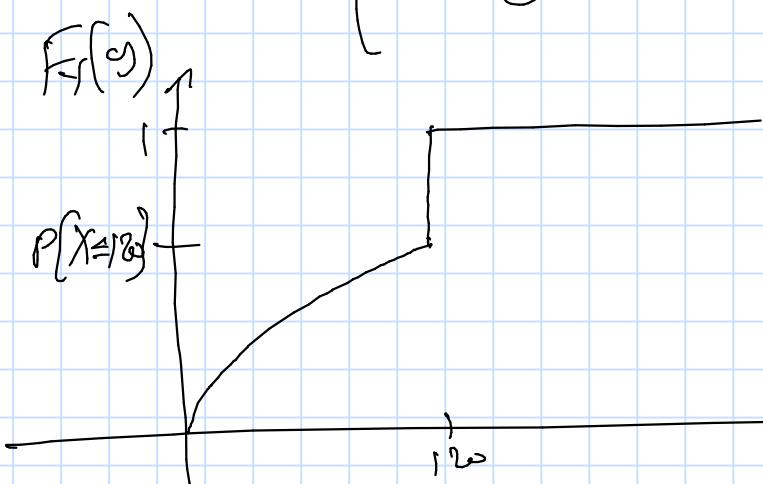
when $X \geq 120$, where, as before, X is the time for expiring of the timer.

Else, the student spends X minutes in the lecture.

Up to
2/5
for an
attempted
resolving.

$$F(y) = P[Y \leq y] = \begin{cases} P[X \leq y] & 0 \leq y < 120 \\ 1. & y \geq 120 \\ 0 & \text{otherwise.} \end{cases}$$

Up to
2/5 for
correct
entries



Up to
3/5 for an
attempted
sketch.