MTH 201: Probability and Statistics

Mid Semester Exam

18/04/2023

Sanjit K. Kaul

No books, notes, or devices are allowed. Just a pen/ pencil and eraser. Any exchange of information related to the quiz with a human or machine will be deemed as cheating. Institute rules will apply. Explain your answers. Show your steps. Don't bother simplifying products and sums that are too time consuming. Leave your answer as products and sums, if you don't have the time. You have about 120 minutes.

Question 1. 40 **marks** An agency is given the task of assessing the knowhow in an institute. The agency decides to do so by asking the professors of the institute a few questions. There are a total of three questions. All three are questions about the full forms of abbreviations. The first question "OBE?" requires the faculty to specify the full form of the abbreviation OBE. The second question "TOC?" requires specifying the full form of the abbreviation TOC and the third question "CO?" queries the full form of CO.

The agency arrives to a room full of professors. It chooses a question from a randomizing device that picks "OBE?" 50% of the times, "TOC?" 30% of the times, and "CO?" otherwise. A professor chosen randomly and with replacement from the room is asked to answer the chosen question. The agency continues to choose and ask professors in the said manner till a professor gives the agency an answer it expects.

For now we will assume that the answer given by a professor is independent of answers given by any professor in any attempt. For the question "OBE?", any professor gives the answer "Outcome Based Education" with probability 0.2, "Ordinary Differential Equations" with probability 0.6, and "Don't Know" with probability 0.2. For "TOC?", the answers given are "Terms of Contract" with probability 0.1, "Theory of Computation" with probability 0.8, and "Don't Know" otherwise. For "CO?", the answers given are "Course Outcomes" with probability 0.2, "Computer Organisation" with probability 0.6, and "Don't Know" with probability 0.2.

Unknown to the professors, the agency expects the answer "Outcome Based Education" to the question "OBE?", the answer "Terms of Contract" for "TOC", and the answer "Course Outcomes" for "CO".

(a) (1 marks) What is the probability $P[\text{OBE?}]$ that the question "OBE?" is asked?

(b) (5 marks) Suppose "OBE?" was asked. Derive the conditional distribution (PMF) of the number of professors the agency will ask the question.

(c) (10 marks) Derive the distribution (PMF) of the number of professors the agency will ask any question chosen by the randomizing device.

(d) (5 marks) Suppose it is known that "OBE?" was **not** asked. Derive the conditional distribution (PMF) of the number of professors the agency will ask the question. [*Hint: The key step is to calculate the probabilities of each of the other two questions, given that "OBE?" was* **not** *asked. Everything else, you may have calculated in a part above.*]

(e) (5 marks) Suppose it is known that the agency asked ≤ 10 professors. Derive the conditional PMF of the total number of professors the agency asks the question.

For all parts that follow, it is known that the question "OBE?" was asked. As before, the agency stops asking as soon as it gets the answer it is looking for. However, it asks no more than two professors. Also, for each of "Outcome Based Education" and "Ordinary Differential Equations", a professor can't give the answers if they were given earlier. That is the professor must choose from the set of other answers (with probabilities of choosing the answers appropriately updated). A professor can however answer "Don't Know" irrespective of whether the same was given as an answer earlier.

(aa) (7 marks) Derive the distribution of the number of professors the agency asks the question.

(bb) (7 marks) RV X takes a value 1 in case the agency gets the answer it expects. Else X takes the value 0. Derive $E[X]$ and $\text{Var}[X]$.

(Q1) Three questions OBE?, TOC?, CO?

$P[OBE?] = 0.5 \quad P[TOC?] = 0.3 \quad P[CO?] = 0.2$

$P[\text{Answer is "Outcome Based Education"} \mid OBE?] = 0.2$

$P[\text{Answer is "Ordinary Diff Eq"} \mid OBE?] = 0.6$

$P[\text{Answer is Don't know} \mid OBE?] = 0.2$

$P[\text{Answer is Terms of Contract} \mid TOC?] = 0.1$

$P[\text{Answer is Theory of Computation} \mid TOC?] = 0.8$

$P[\text{Answer is Don't know} \mid TOC?] = 0.1$

$P[\text{Answer is Course Outcomes} \mid CO?] = 0.2$

$P[\text{Answer is Computer Org} \mid CO?] = 0.6$

$P[\text{Answer is Don't know} \mid CO?] = 0.2$

(a) $P[\text{OBE?}] = 0.5.$ 1 if correct.

(b) Given OBE? is asked. Let N be the no. of prof. The agency asks the professors till it gets the correct answer, which is Outcome Based Education.

Given OBE?, the conditional PMF of the number of professors asked is

$$\underbrace{P[N=n \mid \text{OBE?}]}_{\text{Identifying the correct probability } 2/5} = \underbrace{\begin{cases} (1-p)^{n-1}p & n=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}}_{\text{Providing the correct distribution } 3/5}$$

where

$$p = P[\text{Answer is Outcome Based Education} \mid \text{OBE?}]$$

$= 0.2.$

The conditional PMF is geometric (0.2).

(c) We no longer condition on a specific question. We want the unconditional distribution of N.

The event

$\{N=n\} = \{N=n, OBE?\} \cup \{N=n, CO?\} \cup \{N=n, TOC?\}$

This is because $\{OBE?, CO?, TOC?\}$ is an event space.

$P[N=n] = P[N=n, OBE?] + P[N=n, CO?] + P[N=n, TOC?]$

$= P[N=n | OBE?]\,P[OBE?]$

$+ P[N=n | CO?]\,P[CO?]$

$+ P[N=n | TOC?]\,P[TOC?].$

This statement or equivalent

5/10

$$P[N=n \mid OBE?] = \begin{cases} (0.8)^{n-1}(0.2) & n=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

$$P[N=n \mid CO?] = \begin{cases} (0.8)^{n-1}(0.2) & n=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

$$P[N=n \mid TOC?] = \begin{cases} (0.9)^{n-1}(0.1) & n=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

1 mark for each correctly stated conditional PMF, in case the $P[N=n]$ is incorrectly calculated.

$$P[N=n] = \begin{cases} [(0.8)^{n-1}(0.2)](0.5) + [(0.8)^{n-1}(0.2)](0.2) + [(0.9)^{n-1}(0.1)](0.3) & n=1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

5/10

(d) Given that OBE? was not asked. The PMF of interest is

Let event $A = \{CO?, TOC?\}$. We are given A. We want $P[N=n \mid A]$, $n \in \mathbb{R}$.

Since either CO? or TOC? must have been asked, the conditional PMF(s) of interest are

$$P[N=n \mid A] = P[N=n, CO? \mid A] + P[N=n, TOC? \mid A]$$

$$= P[N=n \mid CO?, A]\, P[CO? \mid A] + P[N=n \mid TOC?, A)\, P[TOC? \mid A]$$

$$= P[N=n \mid CO?]\, P[CO? \mid A] + P[N=n \mid TOC?] + P[TOC? \mid A]$$

We know $P[N=n \mid CO?]$ & $P[N=n \mid TOC?]$

$$P[CO? \mid A] = \frac{P[CO?, A]}{P[A]} = \frac{P[CO?]}{P[A]}$$

$$= \frac{P[CO?]}{P[CO?] + P[TOC?]} = \frac{P[CO?]}{0.5}$$

$$= \frac{0.2}{0.5} = 0.4.$$

$$P[TOC? \mid A] = 0.6.$$

An attempt to calculate the revised probabilities $P[TOC? \mid A]$ & $P[CO? \mid A]$: 2/3.

The correct steps 5/5

(e) We want

$$P[N=n \mid N \le 10], \quad n \in \mathbb{R}.$$

Identifying the probabi:

1/5

$$P[N=n \mid N\le 10] = \begin{cases} \dfrac{P[N=n, N\le 10]}{P[N\le 10]} & n=1,2,\dots,10 \\ 0 & \text{otherwise} \end{cases}$$

3/5

$$= \begin{cases} \dfrac{P[N=n]}{P[N\le 10]} & n=1,2,\dots,10 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{P[N=n]}{\sum_{k=1}^{10} P[N=k]} & n=1,2,\dots,10 \\ 0 & \text{otherwise} \end{cases}$$

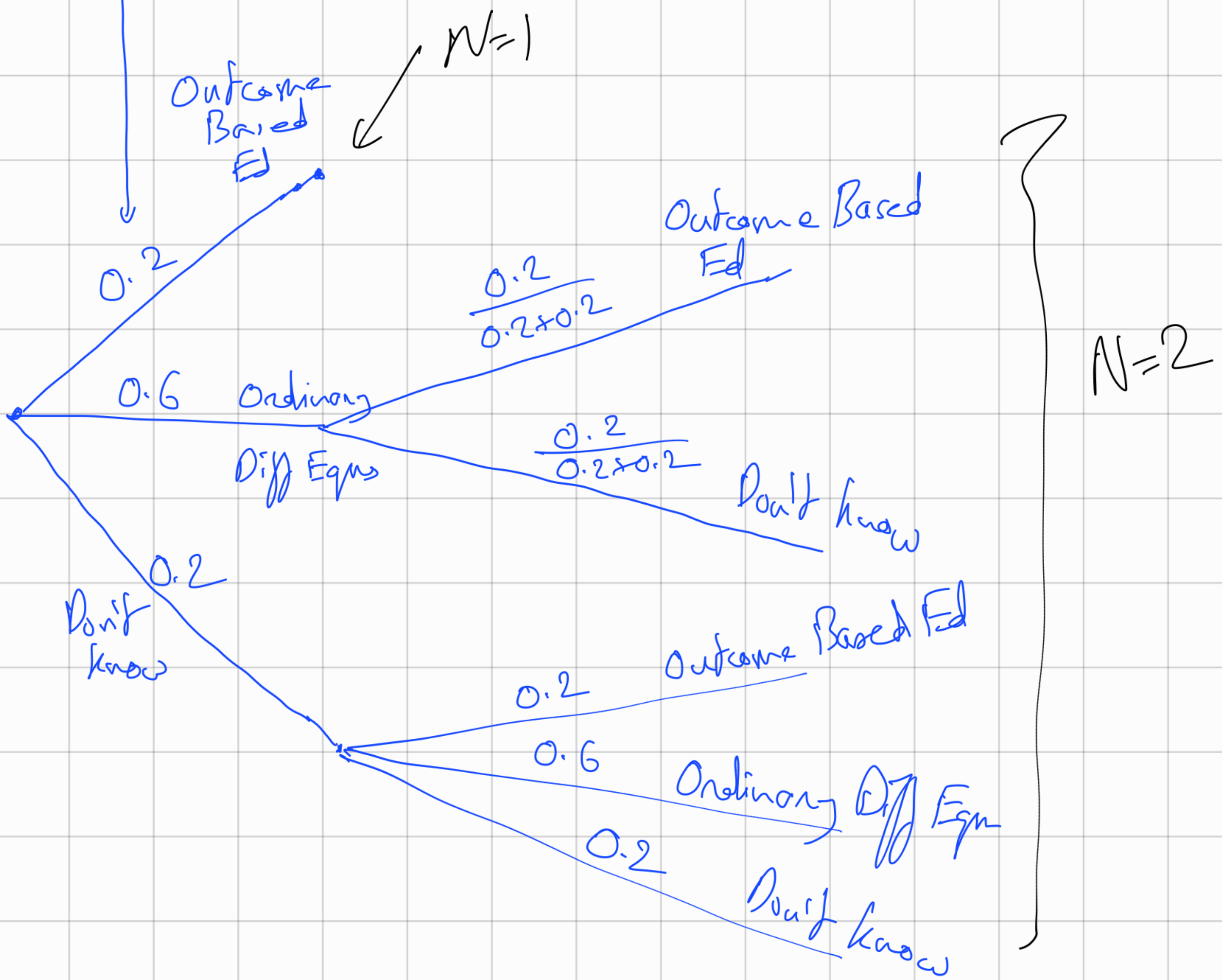
1/5

We calculated $P(N=n)$ earlier.

The tree diagram for $(aa) \Delta (bb)$: (OR equivalent statements of relevant probabilities)

$\left(\frac{8}{(7+7)}\right)$

Answer is $P[\text{Outcome Based Ed} \mid \text{OBE?}]$



(aa) $P[N=n] = \begin{cases} 0.2 & n=1 \\ 0.8 & n=2 \\ 0 & \text{otherwise} \end{cases}$

2/14

(bb) $P[X=1] = P[N=1] + P$[First Prof Answer Ordinary Diff Eqs & 2nd answers Outcome Based Education]

2/14

+ P[First Prof answers don't know & Second prof answers Outcome Based Ed]

$= 0.2 + (0.6)\left(\frac{0.2}{0.2+0.2}\right) + (0.2)(0.2)$

$= 0.2 + (0.6)(0.5) + 0.04 = 0.24 + 0.3 = \underline{0.54}$

$E[X] = (0.54)(1) = 0.54.$

$1/14$

$Var[X] = (0.54)(1)^2 - (0.54)^2$

$= 0.54(1 - 0.54).$

$1/14.$

Question 2. 40 **marks** You go to a casino. The casino has two rooms. In one of the rooms, the large bets room, a player must bet ₹10 every game. In the other room, the small bets room, a player must bet ₹1 every game. In any room, when a player bets ₹x in a game, the player deposits ₹x in a pot and so does the casino. The pot at the beginning of the game, therefore, has ₹$2x$. If the player wins the game, the player keeps the entire pot of ₹$2x$. That is the player gains ₹x. On the other hand, if the player loses the game, the player gets ₹0 from the pot. That is the player gains ₹$(-x)$.

Suppose the probability of winning any game is p. Let RV X be the money that the player gains at the end of playing a game in the small bets room. Let RV Y be the corresponding gains in the large bets room. Derive the following.

(a) (2.5 marks) PMF of X.

(b) (2.5 marks) PMF of Y.

Suppose a player plays $10n$ (n is a positive integer) games in the small bets room. Let X_i, $i = 1, 2, \ldots, 10n$ be $10n$ random variables, where X_i is player gain from the i^{th} game in the room. Let Σ_X be the total gain at the end of playing $10n$ games in the small bets room. Similarly, let Y_i, $i = 1, 2, \ldots, n$ be n random variables, where Y_i is player gain from the i^{th} game in the large bets room. Let Σ_Y be the total gain at the end of playing n games in the large bets room. For both the rooms, the gains in any game are independent of the gains in any other game.

(c1) (5 marks) Write Σ_X in terms of the RV(s) X_i, $i = 1, \ldots, 10n$. Derive $E[\Sigma_X]$. Feel free to use the properties of the E operator.

(c2) (5 marks) As we will see later in the course, since the gains in the games are independent of each other, Var$[\Sigma_X]$ can be written as the sum of variances of X_i, $i = 1, \ldots, 10n$. Derive Var$[\Sigma_X]$.

(c3) (10 marks) Derive $E[\Sigma_Y]$ and Var$[\Sigma_Y]$.

(c4) (2.5 marks) Compare the expected values of Σ_X and Σ_Y. Also, compare their variances. What do you observe?

Suppose in each of n games, a player chooses to play with equal probability in the small bets room and the large bets room. The choice of room in a game is independent of the choices made in other games. Let Z_i be the gain in the i^{th} such game.

(d1) (5 marks) Derive $E[Z_i]$ and Var$[Z_i]$. Use $E[X_i]$, $E[Y_i]$, Var$[X_i]$, and Var$[Y_i]$ to do so.

(d2) (5 marks) Derive $E[\Sigma_Z]$ and Var$[\Sigma_Z]$.

(d3) (2.5 marks) Compare $E[\Sigma_Z]$ and Var$[\Sigma_Z]$ with the expected value and variances of Σ_X and Σ_Y. What do you observe?

Question 3. 10 **marks** The number of people in a conference room is given by RV K, which is Poisson with parameter α. Specifically, the PMF of K is $P[K = k] = \alpha^k e^{-\alpha}/k!$, $k = 0, 1, \ldots$, and $P[K = k] = 0$ otherwise. Let X be a RV that maps the outcomes $1, 2, \ldots$ in range space S_K to 1 and maps the outcome 0 to 0. Calculate the conditional PMF of K, given $X = 1$.

Question 4. 10 **marks** X is a continuous RV with CDF

$$F_X(x) = \begin{cases} 0 & x < 0, \\ x/2 & 0 \le x \le 1, \\ 0.5 & 1 \le x \le 2, \\ 0.5 + 0.25(x-2) & 2 \le x \le 3, \\ 0.75 & 3 \le x \le 4, \\ 0.75 + 0.25(x-4) & 4 \le x \le 5, \\ 1 & x \ge 5. \end{cases}$$

Calculate the following.

(a) $P[X = 2.5]$

(b) $P[X > 4]$

(c) $P[2 \le X \le 3]$

(d) $P[3.5 \le X \le 4.5]$

Q2

(a) $P[X=x] = \begin{cases} p & x=1 \\ 1-p & x=-1 \\ 0 & \text{otherwise} \end{cases}$

2.5

(b) $P[Y=y] = \begin{cases} p & y=10 \\ 1-p & y=-10 \\ 0 & \text{otherwise} \end{cases}$

2.5

(c) $\sum_X = \sum_{i=1}^{10n} X_i$

2/5

$E\left[\sum_X\right] = E\left[\sum_{i=1}^{10n} X_i\right] = \sum_{i=1}^{10n} E[X_i]$

3/5

$$E[X_i] = p(1) + (1-p)\,{-1} = 2p-1$$

$$\therefore E\left[\sum x\right] = 10n(2p-1)$$

(c 2) It is given that

$$\text{Var}\left[\sum x\right] = \sum_{i=1}^{10n} \text{Var}[X_i]$$

4/5

$$\text{Var}[X_i] = E[x_i^2] - \left(E[X_i]\right)^2$$

$$= p(1)^2 + (1-p)(1)^2 - (2p-1)^2$$

3/5

$$= p + 1 - p - (2p-1)^2 = (2-2p)(2p)$$

$$= 4p(1-p).$$

$$\therefore \text{Var}\left[\sum x\right] = 10n\,4p(1-p)$$

$$= 40np(1-p)$$

(c3) $\Sigma_y = \sum_{i=1}^{n} y_i$

$$E\left[\Sigma_y\right] = E\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} E\left[Y_i\right]$$

5/10

$$E\left[Y_i\right] = p(10) + (1-p)(-10) = 20p - 10$$
$$= 10(2p-1)$$

$$\therefore E\left[\Sigma_y\right] = 10n(2p-1)$$

$$\mathrm{Var}\left[\Sigma_y\right] = \sum_{i=1}^{n} \mathrm{Var}\left[Y_i\right]$$

5/10

$$\mathrm{Var}(Y_i) = E\left[Y_i^2\right] - \left(E\left[Y_i\right]\right)^2$$

$$= 100 - 100(2p-1)^2 = 100(2-2p)(2p)$$
$$= 400p(1-p).$$

$$= 400\,p(1-p).$$

$$\therefore \mathrm{Var}\left[\sum y\right] = 400\,np(1-p)$$

(c4) They have the same expected values (of gains). 0.5/2.5

$\sum y$ has a variance that is 10 times that of $\sum x$. 2/2.5

Playing larger bets for the same total bet money of Rs 10n, results in the same average gains. However, with greater uncertainty (variance).

Which room would you pick!?

(d i)

$$E[Z_i] = E[Z_i \mid \text{if game is a small bet}]\, P[\text{Small bet}] + E[Z_i \mid \text{if game is a large bet}]\, P[\text{Large bet}]$$

$$= E[X_i]\, P[\text{Small bet}] + E[Y_i]\, P[\text{Large bet}]$$

$$= 0.5\,(E[X_i] + E[Y_i])$$

2.5/5

$$= 0.5\,(11(2p-1)) = 5.5\,(2p-1)$$

$$\text{Var}[Z_i] = E[(Z_i - \mu_{Z_i})^2]$$

$$= E[(Z_i - \mu_{Z_i})^2 \mid \text{Small bet}]\, P[\text{Small Bet}] + E[(Z_i - \mu_{Z_i})^2 \mid \text{Large Bet}]\, P[\text{Large Bet}]$$

2.5/5

$$= Var[X_i](0.5) + Var[Y_i](0.5)$$

$$= 0.5(4p(1-p) + 400p(1-p))$$

$$= 202p(1-p)$$

(d2)

$$E\left[\sum z\right] = \sum_{i=1}^{n} E[z_i]$$ 2.5/5

$$= 5.5n(2p-1)$$

$$Var\left[\sum z\right] = 202np(1-p)$$ 2.5/5

(d3) $E\left[\sum z\right]$ is smaller than $E\left[\sum x\right]$ & $E\left[\sum y\right]$ 0.5/2.5

$$Var\left[\sum z\right] > Var\left[\sum x\right]$$ 2/2.5

$$\mathrm{Var}\left[\sum z\right] < \mathrm{Var}\left[\sum y\right]$$ 1

Q3 We want

$$P[K = k \mid X = 1], \quad k \in \mathbb{R}.$$ 1

Since $X=1$, K must take values in the set $\{1, 2, \ldots\}$.

$$P[K = k \mid X = 1] = \begin{cases} \dfrac{P[K=k]}{P[X=1]} & k = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

4

$$= \begin{cases} \dfrac{P[K=k]}{\sum_{k=1}^{\infty} P[K=k]} & k=1,2,\dots \\ 0 & \text{otherwise} \end{cases}$$

5/10

$$= \begin{cases} \dfrac{P[K=k]}{1-P[K=0]} & k=1,2,\dots \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \dfrac{\left(\dfrac{\alpha^k e^{-\alpha}}{k!}\right)}{1-e^{-\alpha}} & k=1,2,\dots \\ 0 & \text{otherwise} \end{cases}$$

Q4

(a) $P[X=2.5]=0$. 2.5

Note that X is a continuous RV from its CDF. Also, there isn't a jump at 2.5.

(b) $P[X>4]=$ 2.5

$$1-P[X\leq 4]$$

$$=1-F_X(4)=1-0.75$$

$$=0.25.$$

(c) $P[2\leq X\leq 3]=P[X\leq 3]-P[X\leq 2]$
$+P[X=2]$

$= F_X(3) - F_X(2) + 0$

2.5

$= 0.75 - 0.5 + 0$

$= 0.25$

(d) $F_X(4.5) - F_X(3.5)$

$= (0.75 + 0.25(4.5 - 4)) - 0.75$

2.5

$= 0.25(0.5) = 0.125$