Manipal Academy of Higher Education

1st Semester Master of Engineering (ME)
ME in Artificial Intelligence & Machine Learning,
ME in AIML & BDA

Degree Examination: November 2024
Subject Code: AML 5103
Subject: Applied Probability & Statistics

Sample Final Exam

3 hours, 10 questions, & 100 points. Keep answers short and precise.

1. [10 points] [CO 2, BT 3] Five friends–*Mohit, Aishwarya, Anjali, Ananya,* and *Manthana*–decide to meet for a quick bite at Canara Mall. There are five possible locations in Canara Mall: *(1) Boda Sheera (2) Pizza Hut (3) McDonald's (4) Domino's (5) Chat Shop.* Here is the scenario:

   - Mohit is thrice as likely to go to Boda Sheera as any one of the other locations.
   - Aishwarya is not a fan of pizza and never goes to Pizza Hut or Domino's. Otherwise, she is equally likely to go to any of the other three places.
   - Anjali is equally likely to go to any of the five places.
   - Ananya and Manthana always go together. They will flip a (fair) coin first: if it is heads, then Ananya decides, otherwise Manthana will decide where to go. Ananya always chooses Domino's and Manthana is equally likely to pick any of the five places.

   Finally, these four groups will decide on which location to go to independently.

   (a) Based on the above information, fill the entries of the table below. It would help to fill row-by-row.

Probability that a particular person or group goes to a particular eat-out					
Location / **Person/Group**	*Boda Sheera*	*Pizza Hut*	*McDonald's*	*Domino's*	*Chat Shop*
Mohit	$p_{11} =$	$p_{12} =$	$p_{13} =$	$p_{14} =$	$p_{15} =$
Aishwarya	$p_{21} =$	$p_{22} =$	$p_{23} =$	$p_{24} =$	$p_{25} =$
Anjali	$p_{31} =$	$p_{32} =$	$p_{33} =$	$p_{34} =$	$p_{35} =$
Ananya & Manthana	$p_{41} =$	$p_{42} =$	$p_{43} =$	$p_{44} =$	$p_{45} =$

   (b) Calculate the probability that they all will end up in the same place. If you are unsure about the (i, j)th entry in the table, leave it as p_{ij} and use it in the calculation.

2. [10 points] [CO 1, BT 3] A hospital administrator codes incoming patients suffering gunshot wounds according to whether they have insurance (coding 1 if they do and 0 if they do not) and according to their condition, which is rated as good (g), fair (f), or serious (s). Consider an experiment that consists of the coding of a random patient.
   (a) Give the sample space of this experiment clearly showing the outcomes. Are the outcomes in the sample space equally likely? Explain briefly.
   (b) Let A be the event that the patient is in serious condition. Specify the outcomes in A.
   (c) Let B be the event that the patient is uninsured. Specify the outcomes in B.
   (d) Give all the outcomes in the event (B^c or A). Explain this event in plain English.

3. [10 points] [CO 2, BT 3] You have tracked the performance of the local meteorologist and compiled the following data:

$$\begin{gathered}\text{P(forecast rain, and actual rain)} = 0.45,\\ \text{P(forecast rain, and actual no rain)} = 0.15,\\ \text{P(forecast no rain, and actual no rain)} = 0.35.\end{gathered}$$

   (a) How often does she forecast rain?
   (b) How often does she make a mistake?
   (c) Given that she had forecast rain in last night's broadcast, what is the probability that it did not rain today?

4. [10 points] [CO 2, BT 4] Suppose that an insurance company classifies people who buy medical insurance from them into one of three classes: good, average, and bad risks. As a data scientist for the company, you have access to the following customer data for the calendar year 2022-23:

Class	**% of customers**	**% who had a major health problem**
Bad risk	20%	25%
Average risk	40%	20%
Good risk	40%	10%

   (a) What is the probability that a new customer will meet with a major health problem during 2023-24?
   (b) Suppose that out of 12 health issues in a year, the probability that one becomes a major health problem is 0.1 for good, 0.3 for average, and 0.8 for bad risk customers, respectively. A new customer comes in reporting that they had 2 major health problems out of 12 in the last year. What is the probability that the person is good risk?

5. [10 points] [CO 1, BT 3] You run APS Mobile Company that offers cell phones with three storage capacities: Small (1GB), Medium (4GB), and Large (8GB). Based on your

market research, you believe that the three sizes will be ordered by a potential customer with probabilities 0.2, 0.5, and 0.3, respectively, for each order independently of the other orders. For each of the following, identify the correct random variable with the associated parameters clearly shown. Using the parameters, calculate the expected value of each random variable. Your answer should look like, for example,

$$X \sim \text{Bin}(n = 12, p = 0.2) \text{ and } E[X] = np = 12 \times 0.2.$$

(a) The number of Medium phone orders out of the next 50 phone orders.

(b) The number of phones that will be ordered until the next Small phone is ordered given that the previous three phones ordered were Large.

(c) Suppose that 40 of the previous 100 orders were for Small phones. You randomly choose 25 of those 100 orders and would like to know the number of Large or Medium phone orders in them.

6. [10 points] [CO 1, BT 3] On average 5.5 users sign-up for an on-line social networking site each minute. What is the probability that:

   (a) *more than* 7 users will sign-up for the social networking site in the next minute;

   (b) *at most* 13 users will sign-up for the social networking site in the next 2 minutes?

7. [10 points] [CO 2, BT 4] A typing agency employs 2 typists. The average number of errors per article is 3 when typed by the first typist and 2 when typed by the second typist. Suppose that an article is equally likely to be typed by either typist.

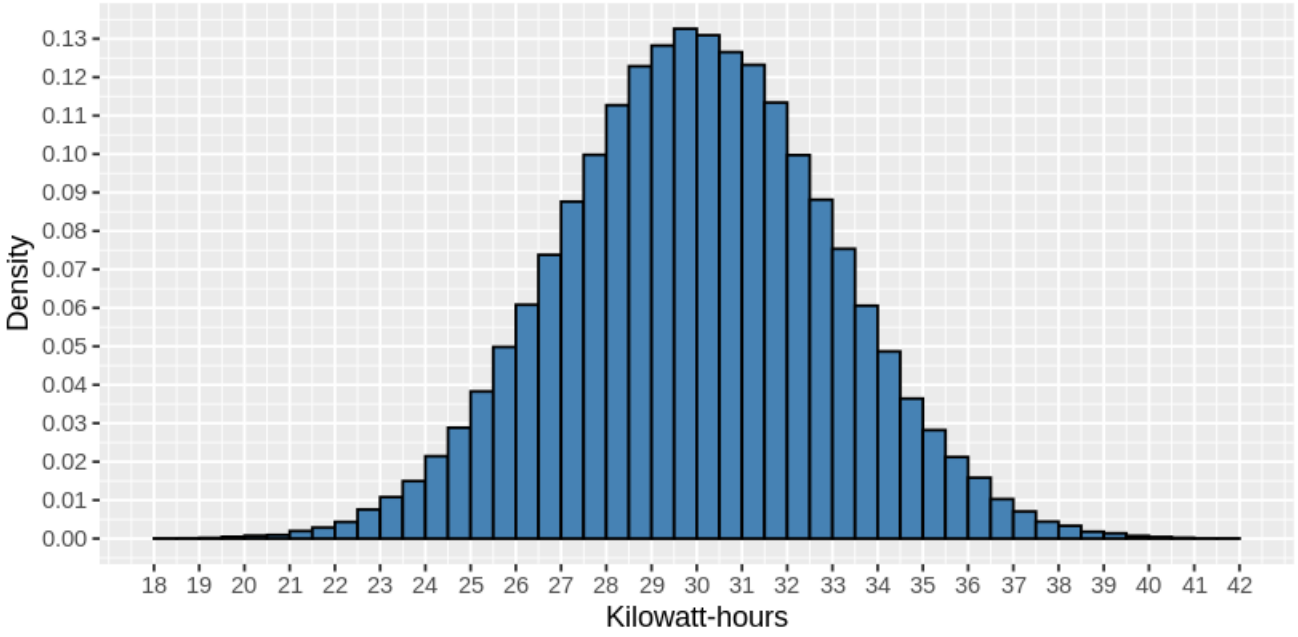
   (a) Let X and Y be the random variables that represent the number of errors made by the first and second typists, respectively. What type of random variables are X and Y, and what are the associated parameters?

   (b) Find the probability that the article will have no errors.

   Hint: *start with $P(\text{no error})$, use the law of total probability followed by Bayes' theorem.*

8. [10 points] [CO 1, BT 3] You know that in your data center, a server fails once every 2000 hours. Suppose a server is typically used from midnight to 10PM every day before switching to a backup. What is the probability that a server will last for more than 2 years without failure?

9. [10 points] [CO 3, BT 4] Consider the following histogram corresponding to a random variable that models the electricity consumed (number of kilowatt-hours) by a machine over a day:



(a) Over the next 1000 days, how many kilowatt-hours on an average per day can we expect the machine to consume?

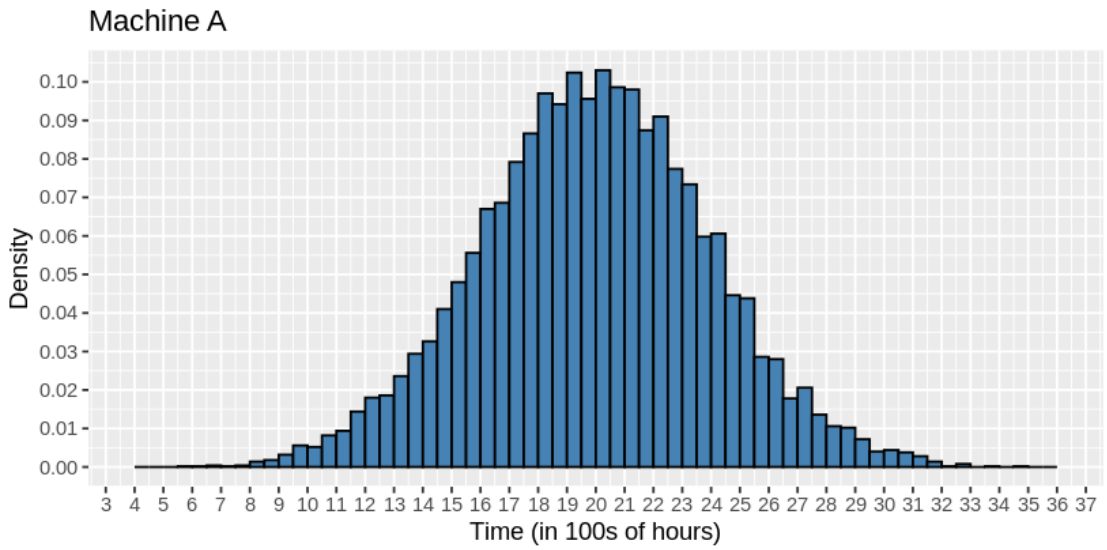
(b) What is more likely to happen in the next day: the machine will consume (i) 38 to 40 kilowatt-hours or (ii) 18 to 20 kilowatt-hours? Explain briefly.

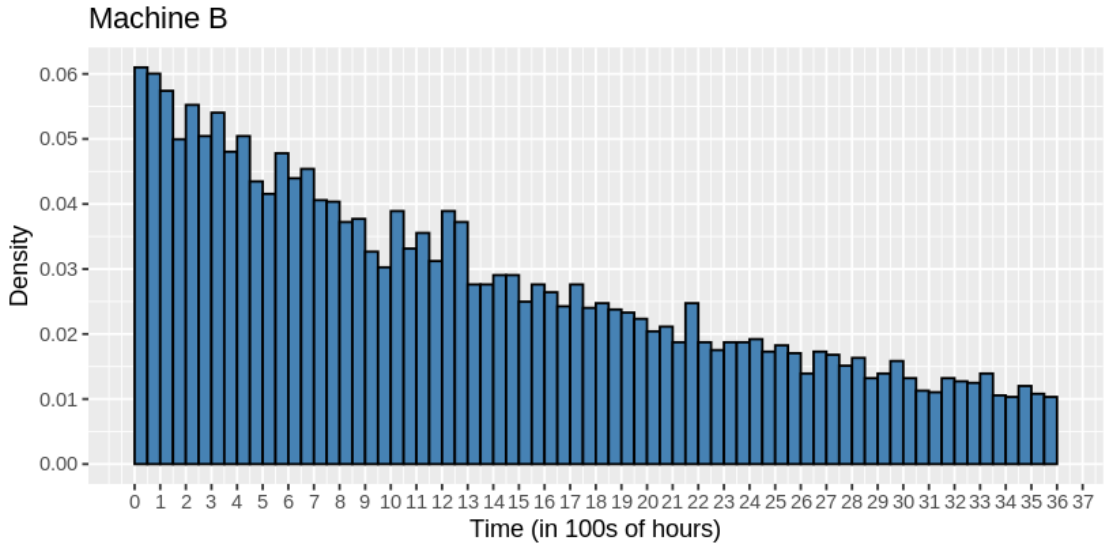
(c) Over the next 1000 days, in how many days, approximately, will the machine consume 24 to 25 kilowatt-hours? Explain briefly.

10. [10 points] [CO 3, BT 4] You run a company called APS Tech that manufactures and sells laptops. APS Tech has outsourced the laptop screen production to two independent companies A and B. The histograms shown below correspond to the life time of laptop screens (in hundreds of hours) that were manufactured by companies A and B. You meet a customer who had bought their laptop from your company, and they are happy with the laptop; they have used the laptop for 18 hundred hours so far with no issues, and the laptop's screen is still working fine at this point in time. You do know that a laptop screen for your company is *twice as likely* to be manufactured by A as it is by B. Explain briefly how you can figure out who among A and B could have manufactured that customer's laptop's screen. You don't have to do any numerical calculation.

Hint: We have $T_A > 18$ and $T_B > 18$. What kind of random variables are T_A and T_B with what parameters?

$$\text{Apply Bayes theorem to } \begin{cases} P(\text{manufactured by } A \,|\, T_A > 18) \\ \text{and} \\ P(\text{manufactured by } B \,|\, T_B > 18). \end{cases}$$


Machine A
Density
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.10
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Time (in 100s of hours)


Machine B
Density
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Time (in 100s of hours)