STATISTICAL METHODS (STAT 3025)

Midterm 1

Total points 90

Instructor : Soumik Banerjee

Instructions

No point will be given for without showing the necessary steps

Each Q carrying 10 *points.Mixing up questions is not allowed.*

Q1: You are given the stem plot below. It's also given that $N = 51$ and leaf unit is 1.

2	37	05
3	38	0
5	39	66
10	40	00267
13	41	007
14	42	0
20	43	113678
(10)	44	0222467799
21	45	0234689
14	46	01247
9	47	0235779
2	48	8
1	49	
1	50	
1	51	4

Find the 5 number summary (MIN,Q1,MEDIAN,Q3 and MAX) of this data using this stem plot. (10 points).

Q2: For any two events A and B with $P(A) > 0$ and $P(B) > 0$. Answer the following questions.

a) Show that $P(A)+P(B)-1 \leq P(A \cup B\ ) \leq P(A)+P(B)$. (Hint: A and B are NOT mutually exclusive) (6 points)

b) Let's say that A and B are independent but NOT mutually exclusive. It is given that $P(A) = 0.3, P(B/A) = 0.4$ Find $P(A \cup B\ )$.(4 points)

(Hint: Recall the expression of $P(A \cup B\ )$ and use the definition of conditional probability under independence)

Q3: For any two events A and B with $P(A) > 0$ and $P(B) > 0$. Answer the following questions.

a) Show that the events $A^{'}$ and B are independent **if it is given that A and B are independent**. (3 points)

b) Show that the events A' and B' are independent **if it is given that A and B are independent**. (3 points)

c) Consider A and B are **NOT INDEPENDENT.** Show that $P(A) = P(A/B)P(B) + P(A/B')P(B')$ (4 points)

(Hint: Recall the theorem of total probability)

Q4: In data scince R and SAS are the two softwares which are most frequently used. In a recent study in a particular public university suggests that 45% of the UG students know R. It has been found that of the students who know R, 70% know SAS and of the students who don't know R, 20% know SAS. Based on this study answer the following questions.

a) Find the probability that any randomly selected student does not know R (3 points)
b) Find the probability that any randomly selected student knows SAS. (3 points)
c) If you know that any student knows SAS find the probability that the student knows R. (4 points)

Q5: Using the information given in **Q4** answer the following questions

a) Find the probability that any student knows R and he **DOES NOT KNOW SAS** (4 points)
b) Find the probability that any student knows either R or SAS or both. (At least one) (4 points)
c) Find the probability that any student knows neither R nor SAS. (2 points)

Q6: Consider two boxes B1 and B2. B1 has 7 red and 5 green balls where B2 has 8 red and 6 green balls. You are selecting one ball at random from B1 (Without replacement) and adding that to B2.Finally selecting one ball from B2. Find the following probabilities

a) What is the probability of selecting a red ball from B1? (3 points)

b) What is the probability of selecting a red ball from B2? (7 points)

Q7: Consider a discrete random variable X with the following CDF.

$$P(X \leq x) = \begin{array}{ll} 0 & x < 0 \\ 0.10 & 0 \leq x < 1 \\ 0.30 & 1 \leq x < 2 \\ 0.50 & 2 \leq x < 3 \\ 0.80 & 3 \leq x < 4 \\ 1 & x \geq 4 \end{array}$$

Based on the information answer the following questions.

a) Find $P(X = 2)$ and $P(X = 4)$ (3 points)

b) Find $P(X > 1)$ and $P(X \leq 4)$ (3 points)
c) $P(1 \leq X \leq 3)$ and $P(2 < X \leq 4)$ (4 points)

Q8:A recent study in an American public University indicates that 30% of the students staying on campus smoke e-cigerate. A newly appointed health officer wants to use this report for her study.She chooses 20 UG resident students at random. Let X be the random variable that denotes how many students smoke e-cigerate out of 20 selected students.

Based on the information answer the following questions.

a) What do you think about the distribution of X? (1 point)
b) What is the probability that the number of smokers is at least 8? (3 points)
c) What is the probability that the number of smokers is less than 5? (3 points)
d) What is the probability that the number of smokers is between 2 and 9 both inclusive? (3 points)

Q9: Suppose the arrivals of patients at a hospital form a Poisson process. The average number of customers per hour is 2.

Let X be the random variable that denotes the number of patients in a given period of 3 consecutive hours.

a) Write explicitly the PMF of X and find the mean and variance of X. (2 points)

b) Find $P(X \geq 1)$ (4 points)

Define a new random variable $Y = 3 + 5X^2$
c) Find $E(Y)$ (4 points)