NATIONAL UNIVERSITY OF SINGAPORE

DEPARTMENT OF STATISTICS AND DATA SCIENCE

ST2334 PROBABILITY AND STATISTICS

MID-SEMESTER TEST SAMPLE PAPER 1

(SEMESTER I, AY 2023/2024)

TIME ALLOWED: 60 MINUTES

Suggested Solutions

INSTRUCTIONS TO STUDENTS

1. Please write your student number only. **Do not write your name.**
2. This assessment contains 15 questions and comprises 7 printed pages.
3. The total marks is 25; marks are equal distributed for all questions.
4. Please answer ALL questions.
5. Calculators may be used.
6. This is an **OPEN BOOK** assessment. Only **HARD COPIES** of materials are allowed.

1. **MULTIPLE RESPONSE: CHOOSE ALL ANSWERS THAT APPLY**

Which of the following can be used as the sample space for the problem: "choose two students from four students to complete a project"? Assume students are labeled as S_1, S_2, S_3, and S_4.

(a) $\Big\{(S_1,S_2),(S_1,S_3),(S_1,S_4),(S_2,S_1),(S_2,S_3),(S_2,S_4),$
$(S_3,S_1),(S_3,S_2),(S_3,S_4),(S_4,S_1),(S_4,S_2),(S_4,S_3)\Big\}$.

(b) $\Big\{\{S_1,S_2\},\{S_1,S_3\},\{S_1,S_4\},\{S_2,S_3\},\{S_2,S_4\},\{S_3,S_4\}\Big\}$.

(c) $\Big\{\{S_1,S_1\},\{S_1,S_2\},\{S_1,S_3\},\{S_1,S_4\},\{S_2,S_2\},\{S_2,S_3\}\Big\}$.

(d) $\Big\{S_1,S_2,S_3,S_4,S_5,S_6\Big\}$

SOLUTION
(a), (b). (a) corresponds to the sample space that order is considered; (b) corresponds to the sample space that order is not considered.

2. **FILL IN THE BLANK**

In how many ways can 3 oaks, 4 pines, and 2 maples be arranged along a property line if one does not distinguish among trees of the same kind?

(Provide your answer in numerical form.)

SOLUTION
Choose 3 places out of 9 slots to plant oaks: $\binom{9}{3}$; and there are 6 slots left. Then choose 4 slots out the rest 6 slots to plant 4 pines; and there are 2 slots left. Then choose 2 slots out the rest 2 slots to plant 2 maples: $\binom{2}{2}$. In total, the number of ways is

$$\binom{9}{3}\cdot\binom{6}{4}\cdot\binom{2}{2} = \frac{9!}{6!3!}\frac{6!}{4!2!}\cdot 1 = \frac{9!}{3!\cdot 4!\cdot 2!} = 1260.$$

3. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

A new Covid test kit detects the virus 90% of the time if a patient is infected. However, it also detects the virus 5% of the time if a patient is uninfected. Given that the overall Covid infection rate is 1%, what is the probability of being infected if your test kit detects the virus?

(a) 0.114
(b) 0.215
(c) 0.154
(d) 0.322

SOLUTION
Let T = Tested Covid; D = Diseased, then

$$P(T|D) = 0.9, \quad P(T|D') = 0.05, \quad P(D) = 0.01.$$

Therefore

$$\begin{aligned} P(D|T) &= \frac{P(T|D)P(D)}{P(T|D)P(D)+P(T|D')P(D')} \\ &= \frac{0.9\times 0.01}{0.9\times 0.01+0.05\times 0.99} = 0.154. \end{aligned}$$

Thus, the answer is (c).

4. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

Suppose $P(F) = P(G) = 0.4$. Which of the following statements must be true?

(a) $P(F \cup G) = 0.8$

(b) $P(F \cup G) = 0.4$

(c) $P(F \cup G) > 0.4$

(d) $P(F \cup G) \leq 0.8$

SOLUTION
Answer: (d).

5. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

There are 10 women and 20 men in a class. Find the number of samples of three that can be formed with two women and one man.

(a) $\binom{30}{3}$

(b) $\binom{30}{1}\cdot\binom{10}{2}$

(c) $\binom{10}{2}\cdot\binom{20}{1}$

(d) $\binom{30}{2}\cdot\binom{20}{1}$

SOLUTION
Answer: (c).

6. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

Player A has entered a golf tournament but it is not certain whether B will enter. Player A has probability 1/6 of winning the tournament if player B enters, and probability 3/4 of winning if player B does not enter the tournament. If the probability that player B enters is 1/3, what is the probability that player A wins the tournament?

(a) 5/9

(b) 7/9

(c) 3/7

(d) 9/11

SOLUTION
(a)

Let $A = \{\text{Player } A \text{ wins the game}\}$, $B = \{\text{Player } B \text{ enters the game}\}$.

From the conditions, we have $P(A|B) = 1/6$, $P(A|B') = 3/4$, $P(B) = 1/3$.

Hence, $P(B') = 1 - P(B) = 2/3$. Applying the law of total probability, we have

$$P(A) = P(A|B')P(B') + P(A|B)P(B) = \frac{3}{4} \cdot \frac{2}{3} + \frac{1}{6} \cdot \frac{1}{3} = 5/9.$$

7. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

   Suppose that A and B are any two events where $P(A) = 0.4$ and $P(A \cap B) = 0.2$. Then $P(A|B) = ?$

   (a) 0.4

   (b) 0.5

   (c) Not enough information to determine

   (d) None of the above

   SOLUTION
   Answer: (c).

8. **TRUE/FALSE**

   Probability density function can not take on values greater than 1.

   - TRUE
   - FALSE

   SOLUTION
   FALSE.

9. **FILL IN THE BLANK**

   Suppose that random variable X has the cumulative distribution function given by

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{100}, & 0 \leq x \leq 10 \\ 1, & x > 10 \end{cases}$$

   Compute $P(X \geq 4)$.

   (Provide your answer in decimal form and round it to two decimal places if necessary.)

   SOLUTION
   $P(X \geq 4) = 1 - P(X < 4) = 1 - F(4) = 1 - 4^2/100 = 0.84.$

10. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

    Let X be a random variable, whose cumulative distribution function is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.2, & 0 \leq x < 2 \\ 0.6, & 2 \leq x < 3 \\ 0.7, & 3 \leq x < 5 \\ 1 & x \geq 5 \end{cases}.$$

    Then $P(1 \leq X < 5) = ?$

(a) 0.1
(b) 0.4
(c) 0.5
(d) 0.8

SOLUTION
(c).
$P(1 \le X < 5) = P(X < 5) - P(X < 1) = F(5-) - F(1-) = 0.7 - 0.2 = 0.5.$

11. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**

The continuous random variable X has the following probability density function

$$f_X(x) = \begin{cases} \frac{1}{8}(1+3x), & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}$$

The median of a continuous random variable Y, denoted by m_Y, is a real number satisfying $P(Y \le m_Y) = 0.5$. What is the median of X?

(a) 4/3
(b) 2/3
(c) 1
(d) 5/3

SOLUTION
For any $x \in [0,2]$,

$$P(X \le x) = \int_0^x \frac{1}{8}(1+3t)dt = \frac{1}{8}\left(x + \frac{3}{2}x^2\right).$$

Set $P(X \le m_X) = 0.5$,

$$\frac{1}{8}\left(m_X + \frac{3}{2}m_X^2\right) = 0.5,$$

which leads to $m_X = 4/3$ or $m_X = -2$ (removed because $m_X \in [0,2]$).

The answer is (a).

12. **FILL IN THE BLANK**

The probability function for random variable X is given by

$$f(x) = \begin{cases} x, & 0 \le x \le 1 \\ 0.5, & 2 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}$$

Compute $E(X)$.

(Provide your answer in decimal form and round it to two decimal places if necessary.)

SOLUTION

$$E(X) = \int_0^1 x \cdot x dx + \int_2^3 x \cdot 0.5 dx = 1/3 + 5/4 = 1.58.$$

13. **FILL IN THE BLANK**

A service station has both self-service and full-service islands. On each island, there is a single regular unleaded pump with two hoses. Let X denote the number of hoses being used on the self-service island at a particular time, and let Y denote the number of hoses on the full-service island in use at that time. The joint probability mass function of X and Y is given in the table below.

x	y		
	0	1	2
0	0.10	0.04	0.02
1	0.08	0.20	0.06
2	0.06	0.14	0.30

Compute $P(X+Y \geq 2)$.

(Provide your answer in decimal form and round it to two decimal places if necessary.)

SOLUTION

0.78

The cells in the table that corresponds to $X+Y \geq 2$ is highlighted in red in the table below.

x	y		
	0	1	2
0	0.10	0.04	0.02
1	0.08	0.20	0.06
2	0.06	0.14	0.30

By adding up these numbers, we obtain $P(X+Y \geq 2) = 0.78$

14. **TRUE/FALSE**

Let $f(x,y)$ be the joint probability function of a random vector (X,Y) (discrete or continuous). If $f_X(1) > 0$, then there must exist a y such that $f(1,y) > 0$.

- TRUE
- FALSE

SOLUTION

TRUE

If otherwise and it is a discrete RV, we have $f_X(1) = \sum_{y \in R_Y} f_X(1,y) = 0$.

If otherwise and it is a continuous RV, we have $f_X(1) = \int_{-\infty}^{\infty} f(1,y)dy = \int_{-\infty}^{\infty} 0dy = 0$.

15. The joint probability function of (X,Y) is given by

$$f(x,y) = \begin{cases} \frac{1}{8}(x+y) & 0 \leq x \leq 2; 0 \leq y \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

Compute $P(Y \geq 1|X=1)$.

(Provide your answer in decimal form and round it to three decimal places if necessary.)

SOLUTION
0.625

The marginal density of X is given by

$$f_X(x) = \int_0^2 \frac{1}{8}(x+y)dy = \frac{1}{8}(2x+2) = \frac{1}{4}(x+1).$$

The conditional probability function

$$f_{Y|X}(y|x=1) = \frac{f(1,y)}{f_X(1)} = \frac{1}{4}(1+y).$$

Therefore

$$P(Y \geq 1|X=1) = \int_1^2 \frac{1}{4}(1+y)dy = \frac{1}{4}\left[y+\frac{y^2}{2}\right]_1^2 = 0.625.$$

END OF PAPER