National University of Singapore

Department of Statistics and Data Science

ST2334 Probability and Statistics

Mid-semester Test

(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 CHOICE: CHOOSE ONE ANSWER ONLY**

Which of the following **MUST** be equal to $A \cup B$? Here, A and B are two events in the sample space S.

(a) $\left[(A \cap B) \cap (A \cup B)\right] \cup (A \cap S)$

(b) $\left[(A \cap B) \cap (A \cup B)\right] \cup (A \cap S)'$

(c) $\left[(A \cap B) \cup (A \cup B)\right] \cup (B \cap S)$

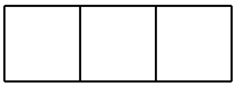
(d) $\left[(A \cap B) \cup (A \cup B)\right] \cup (B \cap S)'$

SOLUTION
(c).

Based on the Venn diagram, (a) is equal to A. (b) is equal to $B \cup A'$. (c) is equal to $A \cup B$. (d) is equal to S.

2. **FILL IN THE BLANK**

Suppose we will use the numbers 1, 3, 5, 76, 125, 876, 987, and 1235 to fill in the three boxes below in such a way that the numbers are arranged in increasing order. For example, if we choose the numbers 987, 125, and 1, we shall put "1, 125, and 987" in the boxes from left to right. How many different ways do we have to fill in the boxes?



(Provide your answer in numerical form.)

SOLUTION
56 (56.00 or 120 are also graded as correct, where the latter is obtained by treating that each number can be used more than once).

We only need to count how many ways to choose three numbers from the 8 numbers, since if the three numbers are given, their order of filling the boxes is deterministic. So we have

$$\binom{8}{3} = \frac{8 \times 7 \times 6}{3!} = 56.$$

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

In a class with 10 boys and 15 girls, a group of 5 students is randomly chosen to work on a project. What is the probability that the group contains 3 boys and 2 girls?

(a) 50/253

(b) 60/253

(c) 70/253

(d) 80/253

SOLUTION

(b).

The total number of choices to select 5 students from the class is $\binom{25}{5}$.

The number of ways to get 3 boys and 2 girls is $\binom{10}{3}\binom{15}{2}$.

Therefore the probability is given by

$$\frac{\binom{10}{3}\binom{15}{2}}{\binom{25}{5}} = 60/253.$$

4. **TRUE/FALSE**

Let A and B be two events in the sample space S. If $P(A) = P(B) = 1$, then A and B **MUST NOT** be independent.

- TRUE
- FALSE

SOLUTION

FALSE

In fact A and B must be independent, since in this case, $P(A \cap B) = 1 = P(A)P(B)$.

5. **FILL IN THE BLANK**

Let A and B be events in the sample space S. Suppose

$$P(A) = 0.6, \quad P(B) = 0.5, \quad P(B|A) = 0.7.$$

Compute $P(A \cap B')$.

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

SOLUTION

0.18.

$P(A \cap B) = P(B|A)P(A) = 0.6 \times 0.7 = 0.42.$

$P(A \cap B') = P(A) - P(A \cap B) = 0.6 - 0.42 = 0.18.$

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

Let A and B be events in the sample space S. Which of the following statements is/are **CORRECT**?

(a) If $P(A) < P(B)$, then we must have $A \neq B$.

(b) If $A \neq B$, then we must have $P(A) \neq P(B)$.

(c) If $A \subset B$, then we must have $P(A) \leq P(B)$.

(d) If $A \cup B = S$, then we must have $P(A) + P(B) \geq 1$.

SOLUTION
(a), (c), and (d).

7. **FILL IN THE BLANK**

Based on a poll, the lecturer of ST2334 learned that 2/3 of his students spent at least 40 hours preparing for the midterm test. For a student who invested at least 40 hours in preparation, there is a 0.9 probability of achieving an "A" grade. Conversely, if the student spent less than 40 hours preparing for the exam, the probability of achieving an "A" grade drops to only 0.3. Given that a student has achieved an "A" grade, what is the probability that the student spent less than 40 hours preparing for the test?

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

SOLUTION
0.14

Let $A = \{\text{achieve } A \text{ grade in midterm}\}$; $B = \{\text{investigate at least 40 hours to study}\}$. Then $P(B) = 2/3$, $P(A|B) = 0.9$, $P(A|B') = 0.3$. The question is asking $P(B'|A)$. We have

$$P(B'|A) = \frac{P(A|B')P(B')}{P(A|B)P(B) + P(A|B')P(B')} = \frac{0.3 \times (1 - 2/3)}{0.9 \times 2/3 + 0.3 \times 1/3} = 1/7 = 0.14286.$$

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

Which of the following is a legitimate probability function?

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

(b) $f(x) = \begin{cases} |x| & \text{for } x = -3, -2, -1, 2, 3 \\ 0 & \text{elsewhere} \end{cases}$

(c) $f(x) = \begin{cases} 0.5x & \text{for } 0 \leq x \leq 2 \\ 0 & \text{elsewhere} \end{cases}$

(d) $f(x) = \begin{cases} 0.5|x| & \text{for } -1 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases}$

SOLUTION
(c)

9. **TRUE/FALSE**

Let $F(x)$ be the cumulative distribution function of a continuous random variable X. Then, for any real number x, it must hold that $F(x) - F(x-) = 0$, where $F(x-)$ is defined as $\lim_{t \uparrow x} F(t)$.

- TRUE
- FALSE

SOLUTION
TRUE

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

Let X be a random variable with $R_X = \{-2, -1, 3, 4\}$. Which of the following is **possibly** the cumulative distribution function of X?

(a) $F(x) = \begin{cases} 0 & x < -2 \\ 0.3 & -2 \le x < -1 \\ 0.8 & -1 \le x < 3 \\ 0.6 & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$

(b) $F(x) = \begin{cases} 0 & x < -2 \\ 0.3 & -2 \le x < -1 \\ 0.4 & -1 \le x < 2 \\ 0.6 & 2 \le x < 4 \\ 1 & x \ge 4 \end{cases}$

(c) $F(x) = \begin{cases} 0 & x \le -2 \\ 0.3 & -2 < x \le -1 \\ 0.4 & -1 < x \le 3 \\ 0.6 & 3 < x \le 4 \\ 1 & x > 4 \end{cases}$

(d) None of the given options

SOLUTION
(d).

11. **FILL IN THE BLANK**

Assume that $F(x)$ below is the cumulative distribution function of a random variable X:

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.4 & 0 \le x < 1 \\ 0.6 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}.$$

Compute $V(X)$, the variance of random variable X.

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

SOLUTION
0.8 (0.80 is also graded as correct).

The probability function is given by $f(x) = 0.4, 0.2, 0.4$ for $x = 0, 1, 2$. We have

$$\begin{aligned} E(X) &= 0.4(0) + 0.2(1) + 0.4(2) = 1. \\ E(X^2) &= 0.4(0)^2 + 0.2(1)^2 + 0.4(2)^2 = 1.8. \end{aligned}$$

Therefore $V(X) = E(X^2) - (E(X))^2 = 1.8 - 1^2 = 0.8.$

12. **FILL IN THE BLANK**

Assume that $f(x)$ below is the probability function for a random variable X:

$$f(x) = \begin{cases} 0.5 & -1 < x \leq 0 \\ x & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}.$$

Compute $E(X)$.

SOLUTION
0.08 (0.083 is also graded as correct).

Based on the structure of $f(x)$, X must be a continuous random variable. We have

$$E(X) = \int_{-1}^{0} x \cdot 0.5dx + \int_{0}^{1} x \cdot xdx = -0.25 + 1/3 = 0.0833$$

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

Let (X,Y) be a discrete random vector, whose joint probability function is given blow:

x	y		
	1	2	3
0	0.02	0.04	0.10
1	0.08	0.10	0.06
2	0.06	0.24	0.30

Then $P(X = Y) = ?$

(a) 0.42

(b) 0.32

(c) 0.24

(d) None of the given options

SOLUTION
(b)

$P(X = Y) = P(X = 1, Y = 1) + P(X = 2, Y = 2) = 0.08 + 0.24 = 0.32.$

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

Let (X,Y) be a discrete random vector, whose joint probability function is given blow:

x	y		
	1	2	3
0	0.02	0.04	0.10
1	0.08	0.10	0.06
2	0.06	0.24	0.30

Then $P(X \geq Y|X = 2) = ?$

(a) 0.5

(b) 0.7

(c) 0.9

(d) None of the given options

SOLUTION

(a).

$P(X \geq Y|X=2) = \dfrac{P(X \geq Y; X=2)}{P(X=2)} = \dfrac{P(Y \leq 2; X=2)}{P(X=2)}$, where we can compute $P(Y \leq 2; X=2) = 0.06 + 0.24 = 0.3$ and $P(X=2) = 0.06 + 0.24 + 0.30 = 0.6$. Therefore $P(X \geq Y|X=2) = 0.3/0.6 = 0.5$.

15. **TRUE/FALSE**

Let (X,Y) be a continuous random vector with the corresponding joint probability function denoted as $f(x,y)$. Then, for any real numbers x and y, it must hold that $f(x,y) \leq f_X(x)$, where $f_X(x)$ represents the marginal probability function of X.

- TRUE
- FALSE

SOLUTION

FALSE

END OF PAPER