NATIONAL UNIVERSITY OF SINGAPORE

DEPARTMENT OF STATISTICS AND DATA SCIENCE

ST2334 PROBABILITY AND STATISTICS

SEMESTER I, AY 2025/2026

Midterm Test (Sample Paper A): Suggested Solutions

- This assessment contains 15 questions.
- The total marks is 30; each question is worth 2 marks.
- Please answer ALL questions.
- Calculators of any kind are allowed.

1. **TRUE/FALSE**

Consider an experiment of **rolling two six-faced regular FAIR dice**. Suppose that the problem of interest is "the numbers that show on the top faces".

If the dice are **not distinguishable**, then the sample space S contains 36 elements.

SOLUTION

The claim is FALSE.

If the dice are **not distinguishable**, we **cannot** tell die 1 apart from die 2. Therefore

- $\{3,4\}$ means that between die 1 and die 2, we have the numbers 3 and 4; and
- $\{4,3\}$ will mean the same.

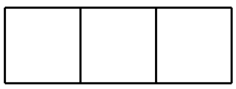
So we have 21 elements in

$$S = \begin{array}{lllllll} \{ & \{1,1\}, & \{1,2\}, & \{1,3\}, & \{1,4\}, & \{1,5\}, & \{1,6\}, \\ & & \{2,2\}, & \{2,3\}, & \{2,4\}, & \{2,5\}, & \{2,6\}, \\ & & & \{3,3\}, & \{3,4\}, & \{3,5\}, & \{3,6\}, \\ & & & & \{4,4\}, & \{4,5\}, & \{4,6\}, \\ & & & & & \{5,5\}, & \{5,6\}, \\ & & & & & & \{6,6\} \quad \} \end{array}$$

Whether the dice are fair or not, we have the same sample space.

2. **FILL IN THE BLANK**

Suppose we will choose THREE numbers from 1, 3, 5, 76, 125, 876, 987, and 1235, with each number being chosen at most once, to fill in the three boxes below such that the numbers are arranged in increasing order from left to right. 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?



ANSWER: ________________. (Leave your answer in numerical form.)

SOLUTION

The answer is 56.

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

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

as our answer.

3. **FILL IN THE BLANK**
A group of 8 friends A,B,C,D,E,F,G,H go to a restaurant. Due to safe-distancing measures, the group needs to split up into two groups of 4. How many ways are there to split the group such that A and B are together but away from C?

ANSWER: ______________. (Leave your answer in numerical form.)

SOLUTION
The answer is 10.

Apart from A, B, C, there are 5 people left. The group with A and B only has two more slots. The two groups are determined if and only if we select 2 more persons out of 5 to join A and B. The rest will go with C. So the number of ways is $\binom{5}{2} = 10$.

4. **TRUE/FALSE**
$P(A \cup B \cup C) = 1 - P(A'|B' \cap C')P(B'|C')P(C')$.

SOLUTION
The claim is TRUE.

Note that

$$\begin{aligned} P(A \cup B \cup C) &= 1 - P[(A \cup B \cup C)'] = 1 - P(A' \cap B' \cap C') \\ &= 1 - P(C')P(B'|C')P(A'|B' \cap C'). \end{aligned}$$

5. **FILL IN THE BLANK**
Find $P(B \cap A)$, if

$$P(A') = 1/2, \quad P(B) = 3/8 \quad \text{and} \quad P(B'|A) = 3/4.$$

ANSWER: ______________. (Round your answer to 3 decimal points, if necessary.)

SOLUTION
The answer is 0.125.

Note that

$$P(B' \cap A) = P(B'|A)P(A) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}.$$

Thus

$$P(B \cap A) = P(A) - P(B' \cap A) = \frac{1}{2} - \frac{3}{8} = \frac{1}{8} = 0.125.$$

6. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**
Suppose A and B are two events where $P(A) = 0.4$ and $P(A \cap B) = 0.2$. What is $P(A|B)$?

(a) 0.4
(b) 0.5
(c) Insufficient information to determine
(d) None of the other options

SOLUTION
The answer is (c).

7. **TRUE/FALSE**
Let A and B be two events in the sample space S. If $P(A) > P(B) = 0.8$, then A and B **ARE NOT** independent.

SOLUTION
The claim is FALSE.

Consider $A = S$. Then no matter what is B, $P(A \cap B) = P(B) = P(A)P(B)$.

8. **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 not infected. Given that the overall Covid infection rate is 1%, what is the probability of being infected if your test kit detects the virus?

Pick the option closest to the answer.

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

SOLUTION
The answer is (c).

Let T = "tested positive"; D = "infected", then

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

Therefore

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

9. **MULTIPLE CHOICE: CHOOSE ONE ANSWER ONLY**
Which of the following is a legitimate probability function?

(a) $f(x) = \begin{cases} x^2/3 & \text{for } 0 \le x \le 2 \\ 0 & \text{elsewhere} \end{cases}$

(c) $f(x) = \begin{cases} x^2/30 & \text{for } x = -3, -2, 1, 2, 3 \\ 0 & \text{elsewhere} \end{cases}$

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

(d) $f(x) = \begin{cases} |x|/11 & \text{for } x = -3, -2, -1, 2, 3 \\ 0 & \text{elsewhere} \end{cases}$

SOLUTION
The answer is (d).

(a) & (d): does not integrate to 1.

(c): does not sum to 1.

10. **Fill in the blank**
Assume that $F(x)$ below is the cumulative distribution function of a random variable X:

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.3, & 0 \le x < 1 \\ 0.7, & 1 \le x < 2 \\ 1, & x \ge 2 \end{cases}.$$

Compute $E(X^2)$.

Answer: ____________. (Round your answer 2 decimal places, if necessary.)

Solution
The answer is 1.6.

The probability function is given by $f(x) = 0.3, 0.4, 0.3$ for $x = 0, 1, 2$. We have

$$E(X^2) = 0.3(0)^2 + 0.4(1)^2 + 0.3(2)^2 = 1.6.$$

11. **Fill in the blank**
The cumulative distribution for a random variable X is given by

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{x^2}{4}, & 0 < x < 2 \\ 1, & x \ge 2 \end{cases}.$$

Compute $V(X)$.

Answer: ____________. (Round your answer 2 decimal places, if necessary.)

Solution
The answer is 0.22.

Based on the structure of the cumulative distribution function, the random variable is a continuous random variable. Thus its probability function is $f(x) = dF(x)/dx = \frac{1}{2}x$ for $0 < x < 2$, and $f(x) = 0$ elsewhere. Thus,

$$\begin{aligned} E(X) &= \int_0^2 x \cdot \frac{1}{2} x \,\mathrm{d}x = \frac{1}{6}x^3\Big|_0^2 = \frac{4}{3} \\ E(X^2) &= \int_0^2 x^2 \cdot \frac{1}{2} x \,\mathrm{d}x = \frac{1}{8}x^4\Big|_0^2 = 2 \end{aligned}$$

and $V(X) = E(X^2) - (E(X))^2 = 2 - \frac{16}{9} = \frac{2}{9} = 0.22$.

12. **True/False**
Let X be a continuous random variable; and let Y be a discrete random variable. It is possible for us to find a real number a, such that $P(X = a) > P(Y = a)$.

Solution
The claim is FALSE.

Since X is continuous, it is always true that $P(X = a) = 0$. However, $P(Y = a) \ge 0$.

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

ANSWER: ______________. (Round your answer to 3 decimal points, if necessary.)

SOLUTION
The answer is 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

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 exists a y such that $f(1,y) > 0$.

SOLUTION
The claim is TRUE.

Otherwise, $f(1,y) = 0$ for all y.

- In the discrete case, we have $f_X(1) = \sum_{y \in R_Y} f(1,y) = 0$.
- In the continuous case, we have $f_X(1) = \int_{-\infty}^{\infty} f(1,y)\,\mathrm{d}y = \int_{-\infty}^{\infty} 0\,\mathrm{d}y = 0$.

Both contradict the given condition that $f_X(1) > 0$.

15. **Fill in the blank**
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 $P(Y \geq 1 | X = 1)$.

Answer: ______________. (Round your answer to 3 decimal points, if necessary.)

Solution
The answer is 0.625.

The marginal density of X is given by

$$f_X(x) = \int_0^2 \tfrac{1}{8}(x+y)\,\mathrm{d}y = \tfrac{1}{8}(2x+2) = \tfrac{1}{4}(x+1).$$

The conditional probability function is

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

Therefore

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

End of Paper