Probability Theory

Tutorial 5

1. The Mathematics Final Exam is multiple choice (MC). It consists of 20 exercises, each with 4 options (only one option is correct). To **pass** the exam, two conditions must be met simultaneously:

   (a) The number of correct answers must be **greater than** the number of incorrect answers.

   (b) There must be a **minimum of 8** correct answers.

   The following piece of advice is common among students: *"Only answer if you know the answer."* Is this true? Of course, if one knew the answers to all questions, the most sensible thing is to answer the whole exam. So, we assume the student does not know the answers and answers all questions randomly and independently.

   Initially, let's compare two scenarios:

   A Answer **15** questions.

   B Answer **20** questions.

   Which of the two scenarios yields a higher probability of passing? Is there a better scenario? What happens if the only restriction is having 8 correct answers?

2. In a cage, there is one mouse and two chicks. The probability that each animal leaves the cage in the next hour is 0.3, and this event is independent for all animals. After one hour, the following random variables are defined:

   - X: The number of legs remaining in the cage. (mouse has 4 legs and the chicks 2 legs)
   - Y: The number of animals remaining in the cage.

   Find the joint probability mass function of (X, Y) and determine the marginal distributions . Are X and Y independent?

3. A retailer verifies that the demand for cars (X) is a random variable with a **Poisson distribution** with parameter $\lambda = 2$ cars per week. The retailer restocks every Monday morning to have 4 cars in stock ($S = 4$).

   The questions posed are:

   (a) What is the probability of selling out the entire stock during the week?

   (b) What is the probability that the retailer is unable to fulfill at least one order?

   (c) What is the minimum stock needed to ensure $P(\text{fulfilling all orders}) \geq 0.99$?

   (d) What is the distribution of the number of cars sold per week?

   (e) If the profit is 1000 € per car sold and the fixed cost is 800 €, what is the distributon of the Weekly Profit (P_G)?

Problem 1

The Mathematics Final Exam is multiple choice (MC). It consists of 20 exercises, each with 4 options (only one option is correct). To **pass** the exam, two conditions must be met simultaneously:

1. The number of correct answers must be **greater than** the number of incorrect answers.
2. There must be a **minimum of 8** correct answers.

The following piece of advice is common among students: *"Only answer if you know the answer."* Is this true? Of course, if one knew the answers to all questions, the most sensible thing is to answer the whole exam. So, we assume the student does not know the answers and answers all questions randomly and independently.

Initially, let's compare two scenarios:

A Answer **15** questions.

B Answer **20** questions.

Which of the two scenarios yields a higher probability of passing? Is there a better scenario? What happens if the only restriction is having 8 correct answers?

Solution

Since each exercise has 4 options and only 1 is correct, the probability of success (answering correctly at random) is:

$$p = P(\text{success}) = \frac{1}{4} = 0.25$$

Strategy A: Answering 15 Questions

Let S_{15} be the number of correctly answered questions out of 15 attempts.

$$S_{15} \sim B(n = 15, p = 1/4)$$

The two passing conditions are:

1. $S_{15} \geq 8$ (*Minimum of 8 correct*)
2. $S_{15} > (\text{Total answered}) - S_{15}$
   $S_{15} > 15 - S_{15} \implies 2S_{15} > 15 \implies S_{15} > 7.5$

The combined passing condition is $\mathbf{S_{15} \geq 8}$ (since S_{15} must be an integer).

$$\begin{aligned}
P(\text{pass in Scenario A}) &= P(\{S_{15} \geq 8\} \cap \{S_{15} > 7.5\}) \\
&= P(S_{15} \geq 8) \\
&= \sum_{k=8}^{15} p_{S_{15}}(k) \\
&= 1 - P(S_{15} \leq 7) \\
&= 1 - F_{S_{15}}(7) \approx \mathbf{0.01729} \quad \text{(low!)}
\end{aligned}$$

Strategy B: Answering 20 Questions

Let S_{20} be the number of correctly answered questions out of 20 attempts.

$$S_{20} \sim B(n = 20, p = 1/4)$$

The two passing conditions are:

1. $S_{20} \geq 8$ (*Minimum of 8 correct*)
2. $S_{20} > 20 - S_{20} \implies 2S_{20} > 20 \implies S_{20} > 10$

The combined passing condition is $\mathbf{S_{20} > 10}$ (or $\mathbf{S_{20} \geq 11}$), as it is the most restrictive requirement.

$$\begin{aligned}
P(\text{pass in Scenario B}) &= P(\{S_{20} \geq 8\} \cap \{S_{20} > 10\}) \\
&= P(S_{20} > 10) \\
&= \sum_{k=11}^{20} p_{S_{20}}(k) \\
&= 1 - P(S_{20} \leq 10) \\
&= 1 - F_{S_{20}}(10) \approx \mathbf{0.003942} \quad \text{(even lower!)}
\end{aligned}$$

Partial Conclusion: The probability of passing with Strategy A ($n = 15$) is ≈ 4.388 times greater than with Strategy B ($n = 20$). The higher probability of passing is achieved by answering **15** questions.

Intermediate Strategies: Is There a Better One?

Let S_n be the number of correctly answered questions out of n attempts, where $8 \leq n \leq 20$.

$$S_n \sim B(n, p = 1/4)$$

The general passing condition is: $S_n \geq 8$ and $S_n > n - S_n \implies S_n > n/2$.

$$\mathbf{P}(\text{pass}) = \mathbf{P}(\mathbf{S_n} > \max\{\mathbf{7}, \mathbf{n/2}\}) = \mathbf{1} - \mathbf{F_{S_n}}(\max\{\mathbf{7}, \lfloor \mathbf{n/2} \rfloor\})$$

Probability of Passing for Different n

The table shows the probability of passing for different values of n:

n	$\max\{7, \lfloor n/2 \rfloor\}$	**Prob. to Pass**
8	7	0.00002
9	7	0.00010
10	7	0.00041
11	7	0.00124
12	7	0.00282
13	7	0.00563
14	7	0.01026
15	**7**	**0.01730**
16	8	0.00754
17	8	0.01237
18	9	0.00543
19	9	0.00885
20	10	0.00394

The highest probability of passing is obtained by answering **15** questions.

Scenario without Incorrect Answer Restrictions

If the only restriction were having a minimum of 8 correct answers ($S_n \geq 8$), and the condition $S_n > n - S_n$ were eliminated:

$$P(\text{pass w/o restr.}) = P(S_n \geq 8) = 1 - F_{S_n}(7)$$

In this case, the probability is strictly increasing with n, making $\mathbf{n} = \mathbf{20}$ the best strategy. However, under the **actual system** with the double restriction, the optimal strategy is $n = 15$.

Problem 2

In a cage, there is one mouse and two chicks. The probability that each animal leaves the cage in the next hour is 0.3, and this event is independent for all animals. After one hour, the following random variables are defined:

- X: The number of legs remaining in the cage.
- Y: The number of animals remaining in the cage.

Find the joint probability mass function of (X, Y), determine the marginal distributions, check for independence, and calculate a conditional probability.

Solution

The cage contains 1 mouse (4 legs) and 2 chicks (2 legs each). The probability that any individual animal **stays** in the cage is $p = 1 - 0.3 = 0.7$. All outcomes are independent.

The total number of animals is 3 (1 mouse, 2 chicks). The maximum number of legs is $4(1)+2(2) = 8$.

The possible values for the random variables X and Y are:

- Y: Total number of animals remaining in the cage. $Y \in \{0, 1, 2, 3\}$.
- X: Total number of legs remaining in the cage. $X \in \{0, 2, 4, 6, 8\}$.

The entries $p_{X,Y}(x, y)$ are calculated based on the independent Bernoulli trials (Mouse stay ~ 0.7; Chick stay ~ 0.7).

Table 1: Joint PMF $p_{X,Y}(x, y)$ and Marginal PMFs

Legs $X\backslash$ Animals Y	**0**	**1**	**2**	**3**	**Marginal $p_X(x)$**
0	0.027	0	0	0	**0.027**
2	0	0.126	0	0	**0.126**
4	0	0.063	0.147	0	**0.210**
6	0	0	0.294	0	**0.294**
8	0	0	0	0.343	**0.343**
Marginal $p_Y(y)$	**0.027**	**0.189**	**0.441**	**0.343**	**1.000**

Marginal Probability Mass Functions

The marginal PMFs are obtained by summing the rows (for $p_X(x)$) and columns (for $p_Y(y)$) in the joint PMF table (Table 1).

Independence Check

X and Y are independent if and only if $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ for all possible pairs (x, y).

Let's test the pair $(X = 8, Y = 3)$:

- **Joint Probability:** $p_{X,Y}(8, 3) = 0.343$
- **Product of Marginals:** $p_X(8) \cdot p_Y(3) = (0.343) \cdot (0.343) \approx 0.1176$

Since $p_{X,Y}(8, 3) \neq p_X(8)p_Y(3)$, the variables **X** and **Y** are **NOT independent**.

Problem 3

A retailer verifies that the demand for cars (X) is a random variable with a **Poisson distribution** with parameter $\lambda = 2$ cars per week. The retailer restocks every Monday morning to have 4 cars in stock ($S = 4$).

The questions posed are:

1. What is the probability of selling out the entire stock during the week?
2. What is the probability that the retailer is unable to fulfill at least one order?
3. What is the minimum stock needed to ensure $P(\text{fulfilling all orders}) \geq 0.99$?
4. What is the distribution of the number of cars sold per week?
5. If the profit is 1000 € per car sold and the fixed cost is 800 €, what is the distributon of the Weekly Profit (P_G)?

Solution

Let X be the weekly demand. $X \sim \text{Poisson}(2)$. The PMF is $P(X = k) = \frac{e^{-2}2^k}{k!}$. Stock $S = 4$.

Probability of Selling Out $P(X \geq 4)$

$$P(X \geq 4) = 1 - P(X \leq 3)$$

$$P(X \leq 3) = e^{-2} \sum_{k=0}^{3} \frac{2^k}{k!} = e^{-2} \left(1 + 2 + 2 + \frac{4}{3}\right) = \frac{19}{3} e^{-2} \approx 0.8571$$

$$P(X \geq 4) = 1 - 0.8571 \approx \mathbf{0.1429}$$

Probability of Unable to Fulfill ($P(X > 4)$)

$$P(X > 4) = 1 - P(X \leq 4)$$

$$P(X = 4) = \frac{e^{-2}2^4}{4!} = \frac{2}{3} e^{-2} \approx 0.0902$$

$$P(X > 4) \approx 1 - (0.8571 + 0.0902) = 1 - 0.9473 \approx \mathbf{0.0527}$$

Minimum Stock (S) for $P(X \leq S) \geq 0.99$

Checking the CDF for X:

- $P(X \leq 5) \approx 0.9834$
- $P(X \leq 6) \approx 0.9955$

The minimum stock is **6** cars.

Distribution of cars Sold (V)

V is the number of cars sold, defined as $V = \min(X, 4)$. V is a **censored Poisson variable**. Its PMF is:

$$P(V = v) = \begin{cases} P(X = v) & \text{for } v \in \{0, 1, 2, 3\} \\ P(X \geq 4) & \text{for } v = 4 \\ 0 & \text{otherwise} \end{cases}$$

Probability Mass Function of the Profit ($P(G = g)$)

Since $G = 1000V - 800$ implies $V = \frac{g+800}{1000}$, the possible values for G are $g \in \{-800, 200, 1200, 2200, 3200\}$.

The PMF of G is:

$$P(G = g) = \begin{cases} P(X = k) = \frac{e^{-2}2^k}{k!} & \text{if } g = 1000k - 800, \text{ for } k \in \{0, 1, 2, 3\} \\ P(X \geq 4) = 1 - \frac{19}{3}e^{-2} & \text{if } g = 3200 \quad (k = 4) \\ 0 & \text{otherwise} \end{cases}$$