

Question 1

Case Study: A company is analyzing customer purchase amounts over a year. They want to determine if these amounts follow a uniform distribution. How can they identify this?

- A) By checking if the data forms a bell curve when graphed.
- B) By ensuring all outcomes have equal frequency across their range.
- C) By observing clustering at certain values in the dataset.
- D) By calculating skewness and kurtosis values.

Correct Answer: B

Explanation: Uniform distributions are identified by their property that all outcomes occur with equal frequency across their defined range.

Question 2

Scenario: A manufacturing plant experiences an average of 3 equipment failures per hour. What is the probability of observing exactly 5 failures in an hour? Round your answer to the nearest whole percentage.

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

where:

- **x** = Actual number of occurrences
- **e** = Euler's number (≈ 2.718)

Correct Answer: 10

Explanation: For a Poisson distribution with $\lambda = 3$, the probability of exactly 5 events is given by $P(X = 5) = (e^{-3} \times 3^5) / 5! \approx 0.1008$, which rounds to 10%.

Question 3

Case Study: A marketing firm conducts a survey with 15 questions, each having a probability of being answered correctly as 0.4. What is the probability of exactly 6 correct answers? Round your answer to the nearest whole percentage.

$$P(X = r) = \binom{n}{r} p^r (1 - p)^{n-r}$$

where:

- **n** = Number of trials
- **r** = Number of successes
- **p** = Probability of success in a single trial

- $$\binom{n}{r}$$

= Binomial coefficient (number of ways to choose **r** successes in **n** trials), given by:

$$\frac{n!}{r!(n - r)!}$$

Correct Answer: 21

Explanation: The probability is given by the binomial probability formula: $C(15,6) * (0.4)^6 * 0.6^9 \approx 0.207$, which rounds to 21%.

Question 4

Scenario: A researcher conducts a study to determine if a new medication is effective. The null hypothesis (H_0) states that the medication has no effect. If the researcher fails to reject H_0 , what does this imply?

- A) There's strong evidence against "medication has no effect".
- B) H_0 must be true.
- C) There's insufficient evidence to support "medication has any effect".
- D) H_0 has been proven correct.

Correct Answer: C

Explanation: Failing to reject H_0 indicates that there isn't enough evidence to support the alternative hypothesis (H_1).

Question 5

A rare genetic disorder occurs in 1 out of every 10,000 people. In a city of 1,000,000 people, what is the expected number of people with this disorder?

Expected no. of people = $n \times p$,

where

- n is no. of trials(no. of people in city)
- p = prob. of success

Correct Answer: 100

Explanation:

In this genetic disorder case:

- There are only two outcomes (has disorder/doesn't have disorder)
- Each trial (person) is independent - The probability remains constant
- There are a fixed number of trials ($n = 1,000,000$)

These conditions are for a binomial distribution.

Therefore, we will use a binomial distribution with n (No. of trials) = 1,000,000 and p (Probability of success or, prob. of rare genetic disorder) = 10^{-4}

The expected value formula is $E(X) = n \times p$

Therefore,

Expected no. of people = $1,000,000 \times (1/10,000) = 100$ people.

This means the expected number of people with the rare genetic disorder in a city of 1,000,000 is 100.

Question 6

What does it mean if data follows a normal distribution?

- A) Data points are uniformly distributed across all values.
- B) Most values cluster around the mean with symmetrical tails on both sides.
- C) There are no outliers present in the data set.
- D) Data can only take integer values.

Correct Answer: B

Explanation: In a normal distribution, most values cluster around the mean, and as you move away from the center, fewer and fewer values occur, creating symmetrical tails.

Question 7

Case Study: A researcher conducts a hypothesis test to determine if a new fertilizer increases crop yield. The p-value obtained is 0.05. What does this p-value represent?

- A) The probability that the null hypothesis is true.
- B) The likelihood of observing data as extreme as what was observed, assuming the null hypothesis is true.
- C) The threshold for rejecting the null hypothesis.
- D) The confidence level of the results.

Correct Answer: B

Explanation: The p-value indicates how likely it is to observe the collected data if the null hypothesis were true; lower p-values suggest stronger evidence against the null hypothesis.

Question 8

Which of the following statements about NumPy are true?

- A) NumPy stands for Numeric Python
- B) NumPy arrays can only store elements of the same data type

- C) NumPy is slower than standard Python lists for large datasets
- D) NumPy provides powerful tools for working with arrays

Correct Answers: A, B, D

Explanation: NumPy stands for Numeric Python, its arrays store elements of the same data type, and it provides powerful tools for working with arrays. It is faster than standard Python lists for large datasets, not slower.

Question 9

Which of the following NumPy functions can be used to create arrays?

- A) `np.array()`
- B) `np.arange()`
- C) `np.ones()`
- D) `np.zeros()`

Correct Answers: A, B, C, D

Explanation: All of these functions can be used to create NumPy arrays. `np.array()` creates an array from a list, `np.arange()` creates an array with a range of numbers, `np.ones()` creates an array filled with ones, and `np.zeros()` creates an array filled with zeros.

Question 10

What will be output of following code?

```
import numpy as np
arr = np.array([[1,3,-1],[2,4,-2]])
print(f"{arr[1,0]} and {arr[0][-1]}")
```

- A) 1 and -1
- B) 2 and 2
- C) 2 and -1
- D) -1 and -2

Correct Answers: C) 2 and -1

Explanation:

1. The expression `arr[1,0]` accesses the element in the second row (index 1) and first column (index 0), which is 2.
2. The expression `arr[0][-1]` accesses the last element of the first row. In Python, -1 as an index refers to the last element. So this gives us -1.
3. The f-string `f"{arr} and {arr[-1]}"` combines these two values, resulting in "2 and -1".

Question 11

You have a NumPy array `arr = np.array([2,3,4,3*3,5//2])` . Which of the following will create a new array that contains only the odd numbers from `arr` ?

- A) `arr % 2 == 0`
- B) `arr[arr % 2 == 0]`
- C) `arr[arr == 2]`
- D) `arr[arr % 2 != 0]`

Correct Answers: D) `arr[arr % 2 != 0]`

Explanation:

- A) `arr % 2 == 0` only creates a boolean mask for even numbers, it doesn't create a new array.
- B) `arr[arr % 2 == 0]` would select even numbers, not odd ones.
- C) `arr[arr == 2]` would only select elements equal to 2, not all odd numbers.
- D) `arr[arr % 2 != 0]` , Let's break down this expression:
 1. `arr % 2` calculates the remainder when each element is divided by 2.
 2. `arr % 2 != 0` creates a boolean mask where True represents odd numbers (remainder not equal to 0) and False represents even numbers.
 3. `arr[arr % 2 != 0]` uses this boolean mask to select only the elements where the condition is True, i.e., the odd numbers.

Question 12

What is the shape of the resulting array after the following operation?

```
import numpy as np
a = np.ones((2, 3, 4))
b = np.ones((3, 4))
result = a + b
```

- A) (2, 3, 4)
- B) (3, 4)
- C) (2, 3)
- D) Error: shapes not aligned

Correct Answer: A

Explanation: b matrix will match the shape of a by broadcasting and then, addition is performed element wise.

Question 13

Which of the following are valid ways to perform matrix multiplication in NumPy?

- A) np.dot(a, b)
- B) a @ b
- C) np.matmul(a, b)
- D) a * b

Correct Answers: A, B, C

Explanation: np.dot(), @, and np.matmul() are valid ways to perform matrix multiplication. a * b performs element-wise multiplication, not matrix multiplication.

Question 14

Which of the following statements about broadcasting in NumPy are true?

- A) It always creates a copy of the smaller array
- B) It can lead to implicit dimension creation
- C) It only works with arrays of the same dimension
- D) It follows a set of rules to determine compatibility

Correct Answers: B, D

Explanation: Broadcasting can lead to implicit dimension creation and follows a set of rules to determine compatibility. It doesn't always create copies and can work with arrays of different dimensions.

Question 15

Given a 2D NumPy array `arr` , which of the following operations will correctly compute the column-wise sum?

- A) `np.sum(arr, axis=0)`
- B) `arr.sum(axis=1)`
- C) `np.sum(arr, axis=1)`
- D) `arr.sum(axis=0)`

Correct Answers: A, D

Explanation:

- `np.sum(arr, axis=0)` and `arr.sum(axis=0)` correctly compute the **sum along columns**.
- `np.sum(arr, axis=1)` and `arr.sum(axis=1)` compute the **sum along rows**, making them incorrect for column-wise sum.