Harsha 🤍
Yeh sab answers proper assignment format me likh rahi hoon — directly copy kar sakti ho.

Question 1: Define the z-statistic and explain its relationship to the standard normal distribution. How is it used in hypothesis testing?

Answer:

The **z-statistic (z-score)** measures how many standard deviations a data point or sample mean is away from the population mean.

Formula:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Where:
• $\bar{x}$= sample mean
• μ= population mean
• σ= population standard deviation
• n= sample size

Relationship with Standard Normal Distribution:

When we standardize a normal variable using the z-formula, it follows a **standard normal distribution** (mean = 0, standard deviation = 1).

Use in Hypothesis Testing:

• Calculate z-score
• Find p-value from z-table
• Compare p-value with significance level (α)
• Reject or fail to reject null hypothesis

Question 2: What is a p-value?

Answer:

A **p-value** is the probability of obtaining results as extreme as the observed results, assuming the null hypothesis is true.

Interpretation:

• Small p-value → Strong evidence against null hypothesis
• Large p-value → Weak evidence against null hypothesis

If p-value = 0.01:

It means there is only a 1% probability that the observed result occurred by chance.
We reject the null hypothesis at 5% significance level.

Question 3: Compare Binomial and Bernoulli distributions.

Answer:

Bernoulli Distribution:

- Describes a single trial
- Two outcomes: Success (1) or Failure (0)
- Parameter: p (probability of success)

Binomial Distribution:

- Describes number of successes in n independent Bernoulli trials
- Parameters: n (number of trials), p (probability of success)

Relation:

Binomial is the sum of multiple Bernoulli trials.

Question 4: Conditions for Binomial Distribution

Answer:

Binomial distribution is used when:

- Fixed number of trials (n)
- Each trial has only two outcomes
- Probability of success (p) is constant
- Trials are independent

Relation to Bernoulli:

A binomial experiment consists of repeated Bernoulli trials.

Question 5: Key properties of Poisson distribution

Answer:

Poisson distribution models the number of events occurring in a fixed interval of time or space.

Properties:

- Events occur independently
- Mean = Variance = λ
- Used for rare events
- Parameter λ represents average rate

Example:

Number of calls received per minute in a call center.

Question 6: Define probability distribution and PDF. Difference between PDF and PMF.

Answer:

Probability Distribution:

It describes how probabilities are assigned to values of a random variable.

Probability Density Function (PDF):

Used for continuous random variables.
Area under curve = 1.

Probability Mass Function (PMF):

Used for discrete random variables.
Gives probability of exact value.

Difference:

PMF → Discrete values
PDF → Continuous values

Question 7: Explain Central Limit Theorem (CLT)

Answer:

The Central Limit Theorem states that:

As sample size increases, the sampling distribution of the sample mean approaches a normal distribution, regardless of the population distribution.

Conditions:

- Sample size n ≥ 30
- Samples are independent

Example:

Even if population is skewed, sample means will form normal distribution for large n.

Question 8: Compare z-scores and t-scores

Answer:

Z-score:

- Used when population standard deviation (σ) is known
- Sample size ≥ 30
- Uses standard normal distribution

T-score:

- Used when population standard deviation is unknown
- Sample size < 30
- Uses t-distribution

Difference:

T-distribution has heavier tails than normal distribution.

Question 9: Calculate z-score and p-value

Given:

Sample mean = 105
Population mean = 100
Standard deviation = 15
Sample size = 25

Step 1: Calculate z-score

$$z = \frac{105 - 100}{15/\sqrt{25}}$$
$$z = \frac{5}{15/5}$$
$$z = \frac{5}{3}$$
$$z \approx 1.67$$

Step 2: Find p-value

For z = 1.67 (two-tailed test)

p-value ≈ 0.095

Decision:

Since p-value (0.095) > 0.05
Fail to reject null hypothesis.

Python Code:

```
import scipy.stats as stats

import numpy as np


sample_mean = 105

population_mean = 100

std_dev = 15

n = 25


z = (sample_mean - population_mean) / (std_dev / np.sqrt(n))
```

```
p_value = 2 * (1 - stats.norm.cdf(abs(z)))


print("Z-score:", z)
print("P-value:", p_value)
```

Question 10: Simulate Binomial Distribution in Python

Given:

n = 10
p = 0.6
Samples = 1000

Expected Mean:

Mean = n × p
Mean = 10 × 0.6 = 6

Expected Variance:

Variance = n × p × (1 − p)
Variance = 10 × 0.6 × 0.4
Variance = 2.4

Python Code:

```
import numpy as np
import matplotlib.pyplot as plt


n = 10
p = 0.6


data = np.random.binomial(n, p, 1000)


print("Sample Mean:", np.mean(data))
print("Sample Variance:", np.var(data))


plt.hist(data, bins=11, density=True)
plt.title("Binomial Distribution (n=10, p=0.6)")
plt.xlabel("Number of Successes")
```

```
plt.ylabel("Frequency")
plt.show()
```

Interpretation:

The simulated mean should be close to 6 and variance close to 2.4, verifying binomial properties.