Question 1: What is a null hypothesis (H_0) and why is it important in hypothesis testing?

Answer: Null Hypothesis (H_0)

Meaning of Null Hypothesis

The null hypothesis (H_0) is a statement that assumes no effect, no difference, or no relationship between variables in a population.
It represents the default assumption that any observed result is due to random chance.

Importance of Null Hypothesis in Hypothesis Testing

The null hypothesis is important because:

- It provides a baseline for comparison
- It helps decide whether observed results are statistically significant
- It allows the use of statistical tests to make objective decisions
- It reduces bias in decision-making

In hypothesis testing, the null hypothesis is either rejected or not rejected based on sample data.

Answer

The null hypothesis (H_0) states that there is no effect or difference and serves as the foundation of hypothesis testing. It is important because it provides a basis for statistical testing and helps determine whether results are due to chance or represent a real effect.

Question 2: What does the significance level (α) represent in hypothesis testing?

Answer:

Significance Level (α)

Meaning of Significance Level

The significance level (α) represents the maximum probability of rejecting the null hypothesis when it is actually true.
In other words, it is the risk of making a Type I error.

Commonly used values of α are 0.05 and 0.01.

Role of α in Hypothesis Testing

The significance level is important because:

- It sets the threshold for deciding whether results are statistically significant
- It helps determine the critical region
- It controls the chance of false positive results

If the p-value is less than or equal to α, the null hypothesis is rejected.

Answer

The significance level (α) indicates the probability of making a Type I error in hypothesis testing. It defines the cutoff point for rejecting the null hypothesis and helps determine whether results are statistically significant.

Question 3: Differentiate between Type I and Type II errors.

Answer:

Difference Between Type I and Type II Errors

Type I Error

- A Type I error occurs when the null hypothesis is rejected even though it is true
- It is also called a false positive
- Its probability is denoted by α (alpha)

Example:
Concluding that a new medicine works when it actually does not.

Type II Error

- A Type II error occurs when the null hypothesis is not rejected even though it is false
- It is also called a false negative
- Its probability is denoted by β (beta)

Example:
Concluding that a new medicine does not work when it actually does.

Key Differences

Type I Error	**Type II Error**
Rejecting a true null hypothesis	Not rejecting a false null hypothesis
False positive	False negative
Probability = α	Probability = β

Answer

A Type I error occurs when a true null hypothesis is rejected, while a Type II error occurs when a false null hypothesis is not rejected.

Question 4: Explain the difference between a one-tailed and two-tailed test. Give an example of each.

Answer:

One-Tailed Test vs Two-Tailed Test

One-Tailed Test

A one-tailed test is used when the alternative hypothesis specifies the direction of the effect (greater than or less than).

It tests whether a parameter is only higher or only lower than a certain value.

Example:
Testing whether the average score of students is greater than 70.

Two-Tailed Test

A two-tailed test is used when the alternative hypothesis does not specify any direction.

It tests whether a parameter is different from a certain value (either higher or lower).

Example:
Testing whether the average score of students is different from 70.

Key Differences

One-Tailed Test	Two-Tailed Test
Direction is specified	No direction specified
Tests only one side of distribution	Tests both sides of distribution
More power in one direction	Detects differences in both directions

Answer

A one-tailed test checks for an effect in one direction, while a two-tailed test checks for an effect in both directions. One-tailed tests are used when the direction of change is known, whereas two-tailed tests are used when the direction is not known.

Question 5: A company claims that the average time to resolve a customer complaint is 10 minutes. A random sample of 9 complaints gives an average time of 12 minutes and a standard deviation of 3 minutes. At $\alpha = 0.05$, test the claim.

Answer:

A company claims that the average time to resolve a customer complaint is 10 minutes.
A random sample of 9 complaints gives:

- Sample mean $\bar{x}$ =12 minutes
- Sample standard deviation s=3 minutes
- Significance level α=0.05

Step 1: State the Hypotheses

- Null Hypothesis (H_0): μ=10 minutes
- Alternative Hypothesis (H_1): $\mu \neq 10$minutes

(This is a two-tailed test.)

Step 2: Select the Test

Since the sample size is small (n = 9) and population standard deviation is unknown, a one-sample t-test is used.

Step 3: Calculate the Test Statistic

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$T = \frac{12-10}{3/\sqrt{9}} = \frac{2}{1} = 2$$

Step 4: Critical Value

- Degrees of freedom =n−1=8

- At $\alpha=0.05$\alpha = 0.05 (two-tailed),
  $t_{critical}=\pm 2.306$

Step 5: Decision

$$|t_{calculated}|=2<2.306$$

So, the null hypothesis is not rejected.

Conclusion

There is insufficient evidence at the 5% significance level to reject the company's claim.

Final Answer

At $\alpha=0.05$, the claim that the average time to resolve a complaint is 10 minutes is not rejected.

Question 6: When should you use a Z-test instead of a t-test?

Answer:

When to Use a Z-test

A Z-test should be used instead of a t-test when:

- The population standard deviation (σ) is known
- The sample size is large ($n \geq 30$)
- The data is normally distributed or the sample size is large enough for the Central Limit Theorem to apply

In such cases, the Z-test gives reliable results.

When a t-test Is Used (for comparison)

A t-test is used when:

- The population standard deviation is unknown
- The sample size is small ($n < 30$)

Answer

A Z-test is used instead of a t-test when the population standard deviation is known and the sample size is large. A t-test is preferred when the population standard deviation is unknown and the sample size is small.

Question 7: The productivity of 6 employees was measured before and after a training program.

Employee	Before	After
1	50	55
2	60	65
3	58	59
4	55	58
5	62	63
6	56	59

At $\alpha = 0.05$, test if the training improved productivity.

Answer:

The productivity of 6 employees was measured before and after a training program.
At α=0.05\alpha = 0.05α=0.05, test whether the training improved productivity.

Given Data

Employee	Before	After	Difference (After − Before)
1	50	55	5
2	60	65	5
3	58	59	1
4	55	58	3
5	62	63	1
6	56	59	3

Number of employees n=6n

Step 1: State the Hypotheses

- Null Hypothesis (H_0): Training has no effect on productivity (mean difference = 0)
- Alternative Hypothesis (H_1): Training improves productivity (mean difference > 0)

(This is a one-tailed paired t-test.)

Step 2: Calculate Mean Difference

Sum of differences =5+5+1+3+1+3=18

$$\bar{d} = \frac{18}{6} = 3$$

Step 3: Calculate Standard Deviation of Differences

Differences: 5, 5, 1, 3, 1, 3

Standard deviation:

$$Sd \approx 1.79$$

Step 4: Calculate Test Statistic

$$T = \frac{\bar{d}}{sd/\sqrt{n}}$$

$$T = \frac{3}{1.79/\sqrt{6}} \approx 4.11$$

Step 5: Decision Rule

- Degrees of freedom $= n - 1 = 5$
- Critical value at $\alpha = 0.05$ (one-tailed) ≈ 2.015

Since:

$$t_{calculated}(4.11) > t_{critical}(2.015$$

Conclusion

The null hypothesis is rejected.

Final Answer

At the 5% significance level, there is sufficient evidence to conclude that the training program improved employee productivity.

Question 8: A company wants to test if product preference is independent of gender.

Gender	Product A	Product B	Total
Male	30	20	50
Female	10	40	50
Total	40	60	100

At α = 0.05, test independence

Answer:

A company wants to test whether product preference is independent of gender.
Test at $\alpha=0.05$

Given Data

Gender	Product A	Product B	Total
Male	30	20	50
Female	10	40	50
Total	40	60	100

Step 1: State the Hypotheses

- Null Hypothesis (H_0): Product preference is independent of gender
- Alternative Hypothesis (H_1): Product preference is not independent of gender

Step 2: Calculate Expected Frequencies

Formula:

$$E=\frac{(Row\ Total)(Column\ Total)}{Grand}$$

Gender	Product A (E)	Product B (E)
Male	$\frac{50*40}{100} = 20$	$\frac{50*60}{100} = 30$
Female	$\frac{50*40}{100} = 20$	$\frac{50*60}{100} = 30$

Step 3: Compute Chi-Square Statistic

$$χ2=\sum\frac{(0-E)^2}{E}$$

$$= \frac{(30-20)^2}{20}+\frac{(20-30)^2}{30}+\frac{(10-20)^2}{20}+\frac{(40-30)^2}{30}$$

=5+3.33+5+3.33=16.66

Step 4: Decision Rule

- Degrees of freedom =(r−1)(c−1)=(2−1)(2−1)=1
- Critical value at α=0.05,df = 1 → 3.84

Since:

x^2calculated(16.66)>x^2critical(3.84)

Conclusion

The null hypothesis is rejected.

Final Answer

At the 5% significance level, there is sufficient evidence to conclude that product preference is not independent of gender.