Hypothesis Testing

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

Answer: Null Hypothesis (H_0)

Definition:
The null hypothesis (H_0) is a statement that assumes no effect, no difference, or no relationship exists in a population. It represents the status quo or a claim that any observed difference is due to random chance.

Why is the Null Hypothesis Important?

1. **Provides a Starting Point**
   Hypothesis testing begins by assuming H_0 is true and then checking whether the data provides enough evidence to reject it.
2. **Basis for Statistical Testing**
   All test statistics, p-values, and confidence levels are calculated under the assumption that H_0 is true.
3. **Controls Decision-Making Errors**
   It helps manage the risk of Type I error (incorrectly rejecting a true null hypothesis).
4. **Ensures Objectivity**
   By testing evidence against H_0, researchers avoid biased conclusions and rely on data-driven decisions.

Example:

- H_0: The average test score of students is equal to 70.
- H_1 (Alternative): The average test score is not equal to 70.

If the data strongly contradict H_0, it is rejected; otherwise, it is not rejected.

Conclusion:

The null hypothesis is crucial because it provides a clear, testable baseline that allows statisticians to objectively evaluate evidence and draw valid conclusions.

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

Answer: Significance Level (α) in Hypothesis Testing

Definition:
The significance level (α) is the maximum probability of making a Type I error, which means rejecting the null hypothesis (H_0) when it is actually true.

What α Represents:

1. **Threshold for Decision-Making**
   α sets the cutoff point for deciding whether the observed results are statistically significant.
2. **Risk of False Positive**
   It represents the researcher's tolerance for error in wrongly concluding that an effect or difference exists.
3. **Comparison with p-value**
   ○ If p-value ≤ α → reject H_0
   ○ If p-value > α → fail to reject H_0
4. **Common Values of α**
   ○ 0.05 (5%) → most commonly used
   ○ 0.01 (1%) → stricter criterion
   ○ 0.10 (10%) → more lenient criterion

Example:

If α = 0.05, it means there is a 5% risk of rejecting a true null hypothesis.

Conclusion:

The significance level (α) defines how strong the evidence must be to reject the null hypothesis and controls the likelihood of making a Type I error in hypothesis testing.

3. Differentiate between Type I and Type II errors.

Answer: Difference Between Type I and Type II Errors:

In hypothesis testing, errors occur when we make an incorrect decision about the **null hypothesis (H_0)**.

Aspect	Type I Error	Type II Error
Definition	Rejecting a **true** null hypothesis	Failing to reject a **false** null hypothesis
Symbol	α (alpha)	β (beta)
Also called	False Positive	False Negative

What goes wrong	Concluding that an effect exists when it actually does not	Concluding that no effect exists when it actually does
Controlled by	Significance level (α)	Sample size, power of the test
Example (Medicine)	Declaring a drug effective when it is not	Declaring a drug ineffective when it actually works

Simple Example:

- **H_0:** The new teaching method has no effect.
  - **Type I Error:** Concluding the method works when it doesn't.
  - **Type II Error:** Concluding the method doesn't work when it does.

Conclusion:

- **Type I error (α)** → rejecting a true H_0
- **Type II error (β)** → not rejecting a false H_0

Both errors are important, and reducing one often increases the risk of the other.

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

Answer: Difference Between One-Tailed and Two-Tailed Tests

In hypothesis testing, one-tailed and two-tailed tests differ based on the direction specified in the alternative hypothesis.

1. One-Tailed Test

Definition:
A one-tailed test checks for an effect in only one direction (either greater than or less than a specified value).

Characteristics:

- The rejection region is entirely in one tail of the distribution.
- Used when the research question is directional.
- More powerful for detecting an effect in the specified direction.

Example:
A teacher believes a new teaching method increases test scores.

- H$_0$: Mean score ≤ 70
- H$_1$: Mean score > 70

Here, only scores greater than 70 are of interest → one-tailed test (right-tailed).

2. Two-Tailed Test

Definition:
A two-tailed test checks for an effect in both directions (either increase or decrease).

Characteristics:

- The rejection region is split between both tails of the distribution.
- Used when the research question is non-directional.
- More conservative than a one-tailed test.

Example:
A researcher wants to know whether a new teaching method changes test scores.

- H$_0$: Mean score = 70
- H$_1$: Mean score ≠ 70

Here, both higher or lower scores matter → two-tailed test.

Conclusion:

- One-tailed test → tests for an effect in one direction only
- Two-tailed test → tests for an effect in both directions

The choice depends on the research question and hypothesis.

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 α = 0.05, test the claim.**

Answer: Hypothesis Testing: Testing the Company's Claim

Given:

- Claimed mean time (μ_0) = 10 minutes
- Sample mean ($\bar{x}$) = 12 minutes
- Sample standard deviation (s) = 3 minutes
- Sample size (n) = 9
- Significance level (α) = 0.05

Since the population standard deviation is unknown and the sample size is small, we use a one-sample t-test.

Step 1: State the Hypotheses

The company claims the mean is 10 minutes.

- Null hypothesis (H₀): μ = 10
- Alternative hypothesis (H₁): μ ≠ 10 (two-tailed test)

Step 2: Test Statistic Formula

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Step 3: Calculate the Test Statistic

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

Step 4: Critical Value

- Degrees of freedom (df) = n − 1 = 8
- At α = 0.05 (two-tailed), critical t-value ≈ ±2.306

Step 5: Decision Rule

- If $|t| > 2.306$, reject H₀
- Here, $|t| = 2 < 2.306$

Step 6: Conclusion

 Fail to reject the null hypothesis

Final Conclusion:

At the 5% significance level, there is not enough statistical evidence to reject the company's claim.

The data does not significantly differ from the claimed average complaint resolution time of 10 minutes.

6.   **When should you use a Z-test instead of a t-test?**

Answer: We should use a Z-test instead of a t-test under the following conditions:

When to Use a Z-Test

1. **Population Standard Deviation is Known**

- ○ If the population standard deviation (σ) is known, a Z-test is appropriate.
- ○ If σ is unknown and replaced by the sample standard deviation, a t-test is used.

2. **Large Sample Size**
   - ○ When the sample size is large (n ≥ 30), the sampling distribution of the mean is approximately normal (by the Central Limit Theorem), so a Z-test can be used even if the population is not normally distributed.

3. **Population is Normally Distributed**
   - ○ If the population is normal and σ is known, a Z-test is suitable for any sample size.

When to Use a t-Test (for comparison)

- Population standard deviation is unknown
- Sample size is small (n < 30)
- Data is approximately normal

Summary Table

Condition	Use Z-test	Use t-test
σ known	✅	❌
σ unknown	❌	✅
Large sample (n ≥ 30)	✅	❌
Small sample (n < 30)	❌	✅

Conclusion:

Use a Z-test when the population standard deviation is known and/or the sample size is large. Otherwise, use a t-test.

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 α = 0.05, test if the training improved productivity.

Answer: Hypothesis Testing: Effect of Training on Productivity

This is a paired (dependent) t-test because productivity is measured before and after training for the same employees.

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

Step 1: State the Hypotheses

Since we want to test whether productivity improved:

- H_0 (Null Hypothesis): $\mu_d = 0$ (No improvement)
- H_1 (Alternative Hypothesis): $\mu_d > 0$ (Productivity improved)

➡️ One-tailed test

Step 2: Calculate Mean and SD of Differences

Differences: [5, 5, 1, 3, 1, 3]

- **Mean difference:**

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

- **Standard deviation of differences:**

$$s_d \approx 1.79$$

Step 3: Test Statistic

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

$$t = \frac{3}{1.79/\sqrt{6}} \approx \frac{3}{0.73} \approx 4.11$$

Step 4: Critical Value

- Degrees of freedom (df) = n − 1 = 5
- At α = 0.05 (one-tailed), critical t ≈ 2.015

Step 5: Decision

- Calculated t = 4.11
- Critical t = 2.015

Since 4.11 > 2.015, we reject H_0.

Final Conclusion:

At the 5% significance level, there is strong statistical evidence that the training program improved employee productivity.

The increase in productivity after training is statistically significant.

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: Chi-Square Test for Independence

We want to test whether product preference is independent of gender.

Step 1: State the Hypotheses

- H_0 (Null Hypothesis): Product preference is independent of gender.
- H_1 (Alternative Hypothesis): Product preference is dependent on gender.

Step 2: Observed Frequencies (O)

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

Step 3: Calculate Expected Frequencies (E)

Formula:

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

- **Male, Product A:** $\frac{50 \times 40}{100} = 20$
- **Male, Product B:** $\frac{50 \times 60}{100} = 30$
- **Female, Product A:** $\frac{50 \times 40}{100} = 20$
- **Female, Product B:** $\frac{50 \times 60}{100} = 30$

Step 4: Compute Chi-Square Statistic

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

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

$$\chi^2 = \frac{100}{20} + \frac{100}{30} + \frac{100}{20} + \frac{100}{30}$$

$$\chi^2 = 5 + 3.33 + 5 + 3.33 = 16.66$$

Step 5: Critical Value

- **Degrees of freedom:**

$$(df) = (r-1)(c-1) = (2-1)(2-1) = 1$$

- At $\alpha = 0.05$ and df = 1
  Critical χ^2 value = 3.84

Step 6: Decision

- Calculated χ^2 = 16.66
- Critical χ^2 = 3.84

Since 16.66 > 3.84, we reject H_0.

Final Conclusion:

At the 5% significance level, there is significant evidence that product preference is NOT independent of gender.

Gender influences product preference.