

1. How do you assess the statistical significance of an insight?

You assess statistical significance by performing **hypothesis testing**, typically using a **p-value**.

- If the p-value < significance level (e.g., 0.05), the result is considered **statistically significant**.
- This means the observed effect is unlikely due to chance.

Example: If you're testing whether a new marketing strategy increased sales, a p-value of 0.03 would indicate it's statistically significant at the 5% level.

2. What is the Central Limit Theorem? Explain it. Why is it important?

The Central Limit Theorem states:

The distribution of sample means approaches a normal distribution as the sample size increases, regardless of the population's distribution.

Why it's important:

- It allows us to use normal distribution-based methods (like confidence intervals and hypothesis tests) on sample data, even if the population isn't normally distributed.
- Foundation for inferential statistics.

3. What is the statistical power?

Statistical power is the probability that a test correctly rejects a false null hypothesis (i.e., detects a real effect).

- Power = $1 - \beta$ (β = probability of Type II error)
- Higher power = higher chance of detecting true effects.

Researchers aim for power ≥ 0.8 (or 80%).

4. How do you control for biases?

You can control for biases using:

- Randomization – randomly assign subjects to groups.
- Blinding – prevent knowledge of group assignment.
- Control groups – establish a baseline.
- Matching – pair subjects with similar characteristics.
- Statistical adjustments – regression techniques to control for confounders.

5. What are confounding variables?

A confounding variable is an external factor that influences both the independent and dependent variables, creating a false association.

Example: If studying ice cream sales and drowning rates, temperature could be a confounder (it affects both).

6. What is A/B testing?

A/B testing is a controlled experiment comparing two versions (A and B) of something (e.g., a webpage) to determine which performs better.

- Group A sees version A, Group B sees version B.
- Use metrics (e.g., click rate) and statistical testing to compare.

7. What are confidence intervals?

A/B testing is a controlled experiment comparing two versions (A and B) of something (e.g., a webpage) to determine which performs better.

- Group A sees version A, Group B sees version B.
- Use metrics (e.g., click rate) and statistical testing to compare.