

Name: Akasha Tigalappanavara  
Net ID: at5854

## ## Practice Questions

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

### Answer:

Assessing the statistical significance of insight is important to determine if the results are due to causation or correlation.

Here's a general process:

**Formulate a Hypothesis:** Start with a null hypothesis ( $H_0$ ) that assumes no effect or no difference between groups, and an alternative hypothesis ( $H_1$ ) that assumes the effect or difference you're interested in.

**Select a Significance Level:** Determine the significance level ( $\alpha$ ) that represents the threshold for accepting or rejecting the null hypothesis. Commonly,  $\alpha$  is set at 5%.

**Choose a Statistical Test:** Based on the data and the hypothesis, select an appropriate statistical test.

**Analyze the Data:** Collect and analyze data using the chosen statistical test to obtain the test statistic and p-value.

### Interpret the Results:

**Test Statistic:** This value indicates how much the data diverges from what is expected under the null hypothesis.

**P-value:** This is the probability of obtaining results as extreme as the observed results if the null hypothesis were true. A lower p-value indicates stronger evidence against the null hypothesis.

**Compare P-value to Alpha:** If the p-value is less than the chosen alpha level, you reject the null hypothesis in favor of the alternative hypothesis. This suggests that the observed results are unlikely to have occurred by chance.

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

### Answer:

The Central Limit Theorem (CLT) is a fundamental concept in statistics that describes the behavior of sample means or sample sums when drawn from a population, irrespective of the population's distribution. It states the following:

"When independent random variables are added together, the sum tends toward a normal distribution, regardless of the original variables' distribution."

Significance: The CLT forms the basis for many statistical methods, such as hypothesis testing and confidence intervals. It allows for making inferences about a population from a sample.

3. What is the statistical power?

**Answer:**

Statistical power refers to the probability that a statistical test will correctly reject a false null hypothesis (i.e., detect a true effect) or, in other words, the likelihood of finding a true effect when it exists.

4. How do you control for biases?

**Answer:**

To control the biases, we need to sample the data randomly.

5. What are confounding variables?

**Answer:**

Confounding variables are extraneous factors that can affect the dependent variable in a research study, leading to a misleading association between the independent and dependent variables.

Example: Examining the relationship between exercise and weight loss. The researchers aim to determine whether increased exercise leads to weight loss. However, they might overlook the potential confounding variable of diet.

6. What is A/B testing?

**Answer:**

A/B testing is a way to compare two different versions of something to see which one works better. Imagine you have two versions of a website, an app, an ad, or even a product label. You show one version to some people and the other to different people, then see which one gets a better response.

7. What are confidence intervals?

**Answer:**

The confidence interval gives us the probability that the population parameter will fall between a set of values for a certain proportion of times. Usually, we use confidence intervals that contain either 95% or 99% of expected observations.