1. How do you control for biases?

To control for biases in research or data analysis, various methods can be employed:

- Randomization: Ensuring that the sample represents the population to prevent selection bias.

- Blinding: Masking the group allocation during analysis to prevent observer biases.

- Matching: Pairing participants with similar characteristics in different groups to control for known confounders.

- Statistical adjustments: Using statistical techniques like regression or ANCOVA to adjust for potential confounding variables.

- Validation: Using different methods or data sources to confirm findings and reduce measurement bias.

2. What are confounding variables?

Confounding variables are factors other than the independent variable that might affect the dependent variable in a study, thereby potentially leading to erroneous conclusions about the relationship between the variables. Identifying and controlling for confounders is crucial to isolate the effect of the independent variable.

3. What is A/B testing?

A/B testing is an experimental procedure used to compare two versions of a variable (usually by comparing a control group against a treatment group) to determine which one performs better in a specific context. It's widely used in marketing, web design, and other fields to test changes in strategies or products.

4. When will you use Welch t-test?

The Welch t-test is used when you want to compare the means of two groups and the two groups do not necessarily have equal variances and/or equal sample sizes. It's an adaptation of the Student's t-test and more reliable when the assumption of equal variances is violated.

5. Testing the claim that the average time on the phone is higher than 6 minutes

You have a hypothesis:

- Null Hypothesis (H0): The average time \( \mu \) is 6 minutes, i.e., \( \mu = 6 \).

- Alternative Hypothesis (H1): The average time \( \mu \) is greater than 6 minutes, i.e., \( \mu > 6 \).

This is a one-tailed test. Let's calculate the test statistic and p-value:

Test Statistic for one-sample t-test:

\[ t = \frac{\bar{x} - \mu\_0}{s / \sqrt{n}} \]

Where:

- \( \bar{x} \) = 6.5 minutes (sample mean)

- \( \mu\_0 \) = 6 minutes (hypothesized mean)

- \( s \) = 1.2 minutes (sample standard deviation)

- \( n \) = 50 (sample size)

Calculation:

\[ t = \frac{6.5 - 6}{1.2 / \sqrt{50}} \]

6. Testing the difference in mean scores between two groups

Here, the hypothesis is:

- Null Hypothesis (H0): There is no difference in the mean scores between the two groups, \( \mu\_A = \mu\_B \).

- Alternative Hypothesis (H1): There is a difference, \( \mu\_A \neq \mu\_B \).

Test Statistic for two-sample t-test (using Welch's correction for unequal variances and unequal sample sizes):

\[ t = \frac{\bar{x}\_A - \bar{x}\_B}{\sqrt{\frac{s\_A^2}{n\_A} + \frac{s\_B^2}{n\_B}}} \]

Where:

- \( \bar{x}\_A \) = 75, \( \bar{x}\_B \) = 78

- \( s\_A \) = 8, \( s\_B \) = 7

- \( n\_A \) = 25, \( n\_B \) = 30

Calculation:

\[ t = \frac{75 - 78}{\sqrt{\frac{8^2}{25} + \frac{7^2}{30}}} \]

I'll calculate the t-values and the corresponding p-values for both questions 5 and 6 to see if we can reject the null hypotheses. Let's proceed with these calculations.

Analysis Results

Question 5: Testing if average time on phone per call is greater than 6 minutes

- Calculated t-statistic: 4.20

- p-value: 0.0000568

Since the p-value is much less than 0.05, we reject the null hypothesis. This provides strong evidence to support the claim that the average time spent on the phone per call is greater than 6 minutes.

Question 6: Testing the difference in mean scores between two groups

- Calculated t-statistic: -1.93

- p-value: 0.0589

Here, the p-value is slightly greater than 0.05, which means we do not reject the null hypothesis at the 0.05 significance level. Thus, we conclude that there is not enough statistical evidence to assert that there is a significant difference in the mean scores of the two groups given the data.