Week 5
Name: Varshitha Akula
Net id : va2437

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

To assess the statistical significance of an insight, you typically use statistical tests. Common methods include hypothesis testing, p-values, and confidence intervals. You compare observed data to a null hypothesis to determine if the insight is likely due to chance or a genuine effect. If the p-value is low (typically < 0.05), the insight is statistically significant, suggesting a real effect.

2. What is the Central Limit Theorem? Explain it. Why is it important?
The Central Limit Theorem (CLT) states that, for a sufficiently large sample size, the distribution of sample means from any population will be approximately normally distributed, regardless of the population's original distribution. It's essential because it allows us to make inferences about a population based on sample means, simplifying statistical analysis.

3. What is the statistical power?

Statistical power measures a test's ability to detect a true effect. It's the probability of correctly rejecting a null hypothesis when it's false (i.e., avoiding a false-negative or Type II error). High statistical power is crucial to ensure the accuracy of statistical tests.

4. How do you control for biases?

Controlling for biases involves implementing methods to minimize the impact of systematic errors in data. Strategies may include randomization, blinding, careful data collection, and statistical techniques like regression analysis.

5. What are confounding variables?

Confounding variables are third variables that affect both the independent and dependent variables in a study. They can lead to incorrect conclusions. To control for confounding, researchers often use techniques like matching, stratification, or statistical methods such as regression analysis.

6. What is A/B testing?

A/B testing (split testing) is a method for comparing two versions (A and B) of a webpage, email, or application to determine which one performs better. It helps assess the impact of changes, new features, or variations on user behavior and informs decision-making based on statistically significant differences.

7. What are confidence intervals?

Confidence intervals are ranges around a sample statistic, such as a mean or proportion, that provide a range of values within which the population parameter is likely to fall. For example, a 95% confidence interval for a mean suggests that if you were to sample multiple times, 95% of the intervals would contain

the true population mean.