Statistics Worksheet 4

1. What is central limit theorem and why is it important?

"The central limit theorem asserts that regardless of the form of the population distribution, the sampling distribution of the sample mean approaches a normal distribution as even the sample size increases." The central limit theorem is significant because it is used to evaluate hypotheses and produce confidence intervals.

2. What is sampling? How many sampling methods do you know?

If the population under study has an endlessly large number of individuals, sampling is the only means to obtain data. There are two kinds of sampling techniques:
- Probability sampling includes random selection, which allows you to draw strong statistical conclusions about the whole group.
- Non-probability sampling entails non-random selection based on practicality or other criteria, making data collection easier.

3. What is the difference between type1 and typeII error?

Type I and Type II error occurs while conducting the hypothesis testing. In type I error, null hypothesis comes out to be false even when it is true. It implies, null hypothesis is rejected even when it is true. Simultaneously, in type II error, null hypothesis comes out to be true when it is false. It implies that null hypothesis is accepted even when it is false.

4. What do you understand by the term Normal distribution?

Normal distribution follows a bell-shaped curve for the distribution. Here there is no difference in mean, median, and mode. All are same. There are no outliers present in the distribution.

5. What is correlation and covariance in statistics?

Correlation defines how change in one variable will influence the other variable, whereas, variance defines how two or more variables differs from each other.

6. Differentiate between univariate, Bivariate, and multivariate analysis?

One variable is analyzed in univariate analysis, while two variables are analyzed in bivariate analysis. In multivariate analysis, more than two variables are analyzed and their relationships are examined.

7. What do you understand by sensitivity and how would you calculate it?

Sensitivity is the frequency with which a test accurately gives a positive response for those who have the ailment being tested for. Sensitivity analysis is a strategy for forecasting the result of a choice if the situation differs from the main expectations.

8. What is hypothesis testing? What is H0 and H1? What is H0 and H1 for two-tail test?

Hypothesis testing is a statistical inference method used to determine whether a hypothesis can be sufficiently supported by the data. The null hypothesis (H0) is a declaration that there is no difference, no correlation, or no impact of therapy. The alternative hypothesis, Ha, is a difference, association, or treatment effect assertion. Until shown differently, H0 is considered to be true. However, Ha is the hypothesis that the researcher wishes to support.

9. What is quantitative data and qualitative data?

Quantitative data are numerical measurements of values or counts. Quantitative data are numerical variables. Qualitative data are sorts of information that can be expressed by a name, symbol, or number code.

10. How to calculate range and interquartile range?

Interquartile range is calculated by subtracting the third quartile from first quartile.

11. What do you understand by bell curve distribution ?

The normal distribution, often known as the bell curve, is a continuous probability distribution that is symmetrical on both sides of the mean, with the right side of the centre being an identical mirror image of the left side.

12. Mention one method to find outliers.

The best way to find outliers are with boxplot.

13. What is p-value in hypothesis testing?

Under the premise that the null hypothesis is true, the p-value in null-hypothesis significance testing is the likelihood of generating test findings at least as severe as the result actually observed

14. What is the Binomial Probability Formula?

The binomial distribution allows us to calculate individual and cumulative probability across a given range.

15. Explain ANOVA and it's applications.

The one-way ANOVA can tell you if there are significant differences in the means of your independent variables. When you know how another independent variable's mean differs from the others, you can start to determine which of them is related to your dependent variable and what is driving that behaviour.