

STATISTICS WORKSHEET-1

10. What do you understand by the term Normal Distribution?

Ans. In normally distribution, there is a constant proportion of data points lying under the curve between the mean and a specific number of standard deviations from the mean. In graph form, normal distribution will appear as a bell curve.

11. How do you handle missing data? What imputation techniques do you recommend?

Ans. when number of missing observations is low and in large number , using mean or median can result in loss of variation in data and it is better to use imputations.

Imputation Techniques

- Complete Case Analysis(CCA):- This is a quite straightforward method of handling the Missing Data, which directly removes the rows that have missing data i.e we consider only those rows where we have complete data i.e data is not missing

12. What is A/B testing?

Ans. It is a way to compare the two versions of a variable to find out which performs better in a controlled environment. It requires analysts to conduct some initial research to understand what is happening and determine what feature needs to be tested.

13. Is mean imputation of missing data acceptable practice?

Ans. The process of replacing null values in a data collection with the data's mean is known as mean imputation.

mean imputation decreases the variance of our data while increasing bias. As a result of the reduced variance, the model is less accurate and the confidence interval is narrower.

14. What is linear regression in statistics?

Ans. In statistics, linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables). A linear regression line has an equation of the form $Y = a + bX$, where X is the explanatory variable and Y is the dependent variable.

15. what are the various branches of statistics?

Ans. Two main branches of statistics are [descriptive statistics](#) and [inferential statistics](#)

Descriptive statistics deals with the presentation and collection of data. For example, a physicist studying turbulence in the laboratory needs the average quantities that vary over small intervals of time. The nature of this problem requires that physical quantities be averaged from a host of data collected through the experiment.

Inferential statistics as the name suggests, involves drawing the right conclusions from the statistical analysis that has been performed using descriptive statistics. In the end, it is the inferences that make studies important and this aspect is dealt with in inferential statistics.