1. a)
2. a)
3. b)
4. d)
5. c)
6. b)
7. b)
8. a)
9. c)
10. Normal Distribution is a probability distribution that is symmetric about the mean. In graphical representation it appears as a Bell Curve. Mean and standard deviation are the parameters of normal distribution. Mean and standard deviation of Standard Normal Distribution are 0 and 1 respectively. It is the most common type of distribution assumed in statistical analysis.
11. Three common strategies to deal with missing values are keep, delete and replace. keep strategy is recommended in case missing values are meaningful because they indicate a pattern. Next option is to delete a variable or observation with too many missing values. The replace or impute option replaces missing values with imputation techniques- Missing values in Continous variables can be replaced with mean or median. For categorical variables, missing values can be replaced with mode that is most frequently occurring value. Regression or Tree based imputation can also be used.
12. A/B testing is basically statistical hypothesis testing. It is an analytical technique that estimates population parameters based on sample statistics. The testing is started by making a claim that is hypothesis. Statistical evidence is gathered to accept or reject the hypothesis. Hypothesis consists of Null and Alternative hypothesis. The alternative hypothesis is what we might hope that A/B test will prove to be true. With the help of sample data, null hypothesis is rejected or accepted.
13. Mean imputation is not always a good solution. It may result in bias. Success depends on the individual variable also. It does not preserve the relationship between variables like correlation.
14. Linear regression is the most commonly used predictive analytical technique to model a continuous target variable. The aim of Linear regression is to estimate the linear relation between target variable(dependent variable) and independent variables (predictors). It fits a linear model with coefficients to minimise the residual sum of squares between the observed targets in the data set and the targets predicted by the linear approximation.
15. The two main branches of Statistics are Descriptive and Inferential Statistics. The branch of statistics that focuses on collecting, summarizing, and presenting a set of data is Descriptive statistics. The branch of statistics that analyzes sample data to draw conclusions about a population is Inferential Statistics.