STATISTICS

Choose the correct option:

1. a) True
2. a) Central Limit Theorem
3. b) Modeling bounded count data
4. c) The square of a standard normal random variable follows what is called chi-squared distribution
5. c) Poisson
6. b) False
7. b) Hypothesis
8. a) 0
9. c) Outliers cannot conform to the regression relationship.

Subjective answer type questions:

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

- The normal distribution is the most widely known and used of all distributions. Because the normal distribution approximates many natural phenomena so well, it has developed into a standard of reference for many probability problems.
- Many things actually are normally distributed or very close to it. for example, height and intelligence are approximately normal distributed, measurement errors also often have a normal distribution.
- The normal distribution easy to work with mathematically.
- In many practical cases, the methods developed using normal theory work quite well even when the distribution is not normal.

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

- Complete Case Analysis(CCA):- This is a quite straightforward method of handling the Missing Data, which directly removes the rows that have missing data. We consider only those rows where we have complete data.
- Arbitrary Value Imputation.
- Frequent Category Imputation.

14. What is linear regression in statistics?

If we want to use a variable x to draw conclusions concerning a variable y:y is called dependent , predictor, or explanatory variable, if the relationship between two variables is linear its can be summarized by a straight line. A straight line can be described by an equation.

y=a+bx

a is called the intercept and **b** the slope of the equation, The slope is the amount by which **y** increases when **x** increases by **1** unit.

12. What is A/B testing?

An AB test is an example of **statistical hypothesis testing**, a process whereby a hypothesis is made about the relationship between two data sets and those data sets are then compared against each other to determine if there is a statistically significant relationship or not.

15. What are the various branches of statistics?