Statistics

Q 1 :- (a) True

Q 2 :- (a) Central Limit Theorem

Q 3 :- (b) Modelling bounded count data

Q 4 :- (d) All of the mentioned

Q 5 :- (c) Poisson

Q 6 :- (b) False

Q 7 :- (b) Hypothesis

Q 8 :- (a) 0

Q 9 :- (c) Outliers cannot confirm regression relationship

Q 10 :- Normal Distribution:- Normal distribution is also known as the Gaussian distribution is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph from, normal distribution will appear as a bell curve.

Q 11 :- Handling Missing Data: 1. Mean or Median imputation. When data is missing at random, we can use list-wise or pai-wise deletion of the missing observations.

2. Multivariate Imputation by chained equation (MICE): MICE assumes that the missing data are Missing at Random (MAR). It imputes data on a variable-by-variable basis by specifying an imputation model per variables.

3. Random Forest. Random forest is a non-parametric imputation method applicable to various variable types that works well with both data missing at random and not missing at random.

I will recommend complete case analysis (CCA) imputation technique, Arbitrary Value Imputation and Frequent Category Imputation.

Q 13 :- True, imputing the mean preserves the mean of the observed data. So, if the data are missing completely at random, the estimate of the mean remains unbiased.

Q 14 :- In statistics, linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables.

Q 15 :- Branches of statistics :- Two Branches of statistics Descriptive Statistics and Inferential Statistics.

Descriptive statistics: the branch of statistics that focuses on collecting, summarizing, and presenting a set of data.

Inferential Statistics: the branch of statistics that analyses sample data to draw conclusions about a population.