STATISTICS ASSIGNMENT

ANSWERS:

Objective type:

1. a) True
2. a) Central limit theorem
3. b) Modelling bounded count data
4. d) All of the mentioned
5. c) Poisson
6. b) False
7. b) Hypothesis
8. a) 0
9. c) Outliers cannot conform to the regression relationship

Subjective type:

10. **Normal Distribution**: The Normal Distribution is a continuous probability distribution that is symmetrical around its mean with most values near the central peak.

11. There is no single method to handle missing values. Before applying any method, it is necessary to understand the type of missing values, then check the datatype and skewness of the missing column, and then decide which method is best for a particular problem.
    Imputation techniques:
    a) Imputation with constant value: It replace the missing data with constant or zero value.
    b) Imputation with statistics: It can be mean, median or mode that is replacing the missing data depending on datatype and its skewness.

12. **A/B testing**: It is a way to compare the two versions of a variable to find out which performs better in a controlled environment.

13. Mean imputation is typically considered terrible practice because of two main reasons:
    a) It ignores features correlation.
    b) 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.

    Although, for low or small samples it can be considered fine.

14. **Linear Regression**: It attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is explanatory variable, and the other is the dependent variable.
15. There are basically four division of statistics:
    a) Mathematical statistics
    b) Statistical methods or functions
    c) Descriptive statistics
    d) Inferential statistics