

STATISTICS WORKSHEET - 1

Answers :

1. A) TRUE
 2. A) Central Limit Theorem
 3. B) Modeling 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.
10. Normal distribution, 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 form, normal distribution will appear as a bell curve. Histograms are created over which we plot the probability distribution curve.
11. The simplest strategy for handling missing data is to remove records that contain a missing value using :
1. Mean or Median Imputation.
 2. Multivariate Imputation by Chained Equations (MICE).
 3. Random Forest.
12. A/B testing is a basic randomized control experiment. It is a way to compare the two versions of a variable to find out which performs better in a controlled environment.
13. No because of these two big problems :
1. Mean imputation does not preserve the relationships among variables.
 2. Mean Imputation Leads to An Underestimate of Standard Errors.
14. In statistics, linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables.
15. The two main branches of statistics are descriptive statistics and inferential statistics.
1. Descriptive statistics deals with the presentation and collection of data.
 2. Inferential statistics, as the name suggests, involves drawing the right conclusions from the statistical analysis that has been performed using descriptive statistics.