Statistics Worksheet

1. a) True
2. a) Central Limit Theorem
3. b) Modeling bounded count data
4. d) All of the mentioned
5. c) Poisson
6. a) True
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.

11. Missing data can be dealt with in a variety of ways. I believe the most common reaction is to ignore it. Choosing to make no decision, on the other hand, indicates that your statistical programme will make the decision for you.Your application will remove things in a listwise sequence most of the time. Depending on why and how much data is gone, listwise deletion may or may not be a good idea.Another common strategy among those who pay attention is imputation. Imputation is the process of substituting an estimate for missing values and analysing the entire data set as if the imputed values were the true observed values.And how would you choose that estimate? The following are some of the most prevalent methods

- Mean imputation
- Substitution
- Hot deck imputation
- Cold deck imputation
- Regression imputation

12. A/B testing—also called split testing or bucket testing—compares the performance of two versions of content to see which one appeals more to visitors/viewers. It tests a control (A) version against a variant (B)

version to measure which one is most successful based on your key metrics.

13. No, mean imputation of missing data acceptable practice .Mean imputation is typically considered terrible practice since it ignores feature correlation. Consider the following scenario: we have a table with age and fitness scores, and an eight-year-old has a missing fitness score.

14. Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

15. The two main branches of statistics are descriptive statistics and inferential statistics. Both of these are employed in scientific analysis of data and both are equally important for the student of statistics.