

# **STATISTICS – WORKSHEET 1**

Name: Poonam Kocheta

Batch: Internship 25

## **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. The normal distribution is a continuous probability distribution that is symmetrical around its mean. It is also known as the Gaussian distribution. Its graph is like bell curved that is equally distributed to both sides.
11. Missing data is the main problem in analysis. There are two ways to handle the missing data.
  1. Deletion
  2. ImputationThe different Imputation techniques are:
  1. Imputation using (Mean/Median) Values
  2. Imputation using frequent (Mode) Values
  3. Imputation using Arbitrary Values
12. The process where decision about relationship is made between data types of two sets and sets against each other determine whether there is a significant statistical relationship or not. It is then predicted that one will perform better than the other. In the two data sets, data sets from both pages are observed and compared to determine if two statistically significant improvement over one.

13. Mean imputation of missing data is not an acceptable practice. If we have a large number of data set then Mean Imputation will effect the result.

14. Linear regression shows the linear relationship between many input variables and one output variables. Input variables are independent variable and output variables are dependent variables.

15. There are two branches of statistics:

1. Descriptive statistics-It deals with the collection of data and its presentation.
2. Inferential statistics- It deals with techniques used for the analysis of data, making estimates and drawing conclusions.