

1. True
2. Central Limit Theorem
3. Modelling bound count data
4. All of the above
5. Poisson
6. False
7. Hypothesis
8. 0
9. Outliers cannot conform to the regression relationship

10. A normal distribution is an arrangement of a data set in which most cluster in the middle of the range .A graphical representation of a normal distribution is always a bell curve. The peak is always in the middle and the curve is always symmetrical in normal distribution. In a normal distribution mean , mode and median are all the same. Normal distribution is used in data analytics.

11. Missing values can be handled by three methods:

A) By deleting the data: we can delete that particular data in the data set which is containing missing values and this method can be used iff the dataset contains millions of data.

B) By creating a model : In this method the particular data which is missing is inferred by training the data set of that row or columns . The missing value will be considered as test data and rest values will be considered as training data. This method can be used iff the dataset is small because it is very time consuming for creating models for each and every missing values.

C) By using statistics method: In this method we use mean, mode and median to find out the missing values present in the dataset. And this is the best method to find out the missing values in the dataset.

12. A/B testing is usually used to compare between two products of two different versions. One version will be given to one set of user and the other version will be given to another set of people and decision given by the user helps us to improve the version in better way so this is also called as user experience research methodology.

13.

14.

15. Branches of statistics:

- *Descriptive statistics: It deals with the organization , analysis , summarization and display of data.

- *Probability Statistics: It deals with change or random phenomena.

- *Inferential Statistics: It deals with drawing the conclusions for a sample .