# Statistics Worksheet 1

1. Bernoulli random variables take (only) the values 1 and 0.

Ans:  A) True

2. Which of the following theorem states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases?

Ans: A) Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?

Ans: B) Modeling bounded count data

4. Point out the correct statement.

Ans: D) All of the mentioned

5. -------------- random variables are used to model rates.

Ans: C) Poisson

6. Usually replacing the standard error by its estimated value does change the CLT.

 Ans: B) False

7. Which of the following testing is concerned with making decisions using data?

Ans: B) Hypothesis

8. Normalized data are centered at_____and have units equal to standard deviations of the original data.

Ans: A) 0

9. Which of the following statement is incorrect with respect to outliers?

Ans: C) Outliers cannot conform to the regression relationship.

10. What do you understand by the term Normal Distribution?

Ans: 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.

11. How do you handle missing data? What imputation techniques do you recommend?

Ans:  When dealing with missing data, data scientists can use two primary methods to solve the error: imputation or the removal of data.

The imputation method develops reasonable guesses for missing data. It's most useful when the percentage of missing data is low.

The other option is to remove data. When dealing with data that is missing at random, related data can be deleted to reduce bias. Removing data may not be the best option if there are not enough observations to result in a reliable analysis.

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.

12. What is A/B testing?

Ans: A/B testing, also known as split testing, refers to a randomized experimentation process wherein two or more versions of a variable (web page, page element, etc.) are shown to different segments of website visitors at the same time to determine which version leaves the maximum impact and drive business metrics.

13: Is mean imputation of missing data acceptable practice?

Ans: In my point of view imputation work only when the percentage of missing data is low. i am not clear about imputation techniques. Whether it will work on heavy data or not.

14: What is linear regression in statistics?

Ans: Regression is a statistical method used in finance, investing, and other disciplines that attempts to determine the strength and character of the relationship between one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

15: What are the various branches of statistics?

Ans: The two main branches of statistics. Both are equally important for the of statistics.

1. Descriptive statistics
2. Inferential statistics