

Statistics Worksheet -1

Q 1 to 9 have only one correct answer. Choose the correct option to

Answer your question.

- 1 Bernoulli random variable take (only) the value 1 and 0.
B) False
- 2 Which of the following theorem states that the distribution of averages of iid variable Properly normalized, become that of a standard normal as the size increases?
A) Central limit theorem.
- 3 Point out the correct statements.
D) All of the mentioned
- 4 Which of the following is incorrect with respect to use of Poisson distribution?
B) Modeling bounded count data.
- 5 Poisson distribution random variables are used to model rates.
C) Poisson distribution.
- 6 Usually replacing the standard error by its estimated value does not change the CLT.
B) False
- 7 Which of the following testing is concerned with making decisions using data?
B) Hypothesis
- 8 Normalized data are centered at 0 and have units equal to standard deviations of the Original data.
- 9 Which of the following statements is incorrect with respect to outliers?
A) Outliers can have varying degrees of influence

Q 10 and Q 15 are subjective answer type question , answer them in your own words briefly.

10 What do you understand by 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 for the mean in graphical form , the normal Distribution appears as a “bell curve”.

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

Ans: Mean, Median and Mode

This is one of the most common methods of imputing values when dealing With missing data . In cases where there are a small number of missing observations, Data scientist can calculate the mean or median of the existing observations and insert Them in place of the missing observations Open_In_New.

12 What is A/B testing?

Ans: A/B testing, also known as split testing, is a marketing experiment wherein you split Your audience to test variations on a campaign and determine which performs better. In other words you can show version A of a piece of marketing content to one half of Your audience and version B to another.

13 Is mean imputation of missing data acceptable practice?

Ans: Mean imputation does not preserve the relationship among variable . True, imputing The mean preserves the mean of the observed data . so if the data are missing Completely at random, the estimate of the mean remains unbiased.

14 What is linear regression in statistics?

Ans: Linear regression analysis is used to predict the value of a variable based on the value of

Another variable.

15 What are the various branches of statistics?

Ans: There are three real branches of the statistics: data collection, descriptive statistics and Inferential statistics.