

STATISTICS WORKSHEET-1

Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

1. Bernoulli random variables take (only) the values 1 and 0.
 - a) True
 - b) False

True outcome in binary

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?
 - a) Central Limit Theorem
 - b) Central Mean Theorem
 - c) Centroid Limit Theorem
 - d) All of the mentioned

Answer - a) Central Limit Theorem

3. Which of the following is incorrect with respect to use of Poisson distribution?
 - a) Modeling event/time data
 - b) Modeling bounded count data
 - c) Modeling contingency tables
 - d) All of the mentioned

Answer - b) Modeling bounded count data

4. Point out the correct statement.
 - a) The exponent of a normally distributed random variables follows what is called the log- normal distribution
 - b) Sums of normally distributed random variables are again normally distributed even if the variables are dependent
 - c) The square of a standard normal random variable follows what is called chi-squared distribution
 - d) All of the mentioned
5. _____ random variables are used to model rates.
 - a) Empirical
 - b) Binomial
 - c) Poisson
 - d) All of the mentioned

Answer is Poisson distribution

6. 10. Usually replacing the standard error by its estimated value does change the CLT.
 - a) True
 - b) False

Answer - doesn't change False

7. 1. Which of the following testing is concerned with making decisions using data?
 - a) Probability
 - b) Hypothesis
 - c) Causal
 - d) None of the mentioned

Answer - b) Hypothesis

- 8 4. Normalized data are centered at _____ and have units equal to standard deviations of the original data.
- a) 0
 - b) 5
 - c) 1
 - d) 10

Answer is 0

9. Which of the following statement is incorrect with respect to outliers?
- a) Outliers can have varying degrees of influence
 - b) Outliers can be the result of spurious or real processes
 - c) Outliers cannot conform to the regression relationship
 - d) None of the mentioned

Answer is Outliers cannot conform to the regression relationship

Q10 and Q15 are subjective answer type questions, Answer them in your own words briefly.

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

normal distribution is a continuous probability distribution that is symmetrical around its mean with most values near the central peak. With no outliers & skewness.

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

Missing data can be handled with in a many of ways. some of the most useful methods. Substitution, Hot deck imputation, Cold deck imputation, Regression imputation, Interpolation and extrapolation etc.

12. What is A/B testing?

A/B testing, is a way to compare two versions of something to figure out which performs better.

13. Is mean imputation of missing data acceptable practice?

No, it's not good practice. this ignores feature correlation.

14. What is linear regression in statistics?

Linear regression analysis is used to predict the value of a variable based on the value of another variable. It is used to predict the value of a variable based on the value of another variable.

15. What are the various branches of statistics?

There are two major branches of statistics are known as **descriptive statistics**, which describes the properties of sample and population data, and **inferential statistics**, which uses those properties to test hypotheses and draw conclusions

