

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 (yes)
 - b) False
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 (yes)
 - b) Central Mean Theorem
 - c) Centroid Limit Theorem
 - d) All of the mentioned
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 (yes)
 - d) All of the mentioned
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 (yes)
 - d) All of the mentioned
6. Usually replacing the standard error by its estimated value does change the CLT.
 - a) True
 - b) False (yes)
7. Which of the following testing is concerned with making decisions using data?
 - a) Probability
 - b) Hypothesis (yes)
 - c) Causal
 - d) None of the mentioned
8. Normalized data are centered at _____ and have units equal to standard deviations of the original data.
 - a) 0 (yes)
 - b) 5
 - c) 1
 - d) 10
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 (yes)
 - d) None of the mentioned

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?
----- **It is often called “bell curve” because it looks like bell . In normal distribution :**
 A) mean = median = mode
 B) symmetry about centre
 C) 50% of the value less than the mean and 50% of the value greater than the mean .
11. How do you handle missing data? What imputation techniques do you recommend?
----- **We handle missing data :**
 A) deletion technique
 B) dropping method
 C) Through mean , median and mode .
 I recommend KNN .
12. What is A/B testing?
----- **A/B testing is a basic randomized control experiment. It is a way to compare the two versions of a variable to find out which performs better in a controlled environment.**
13. Is mean imputation of missing data acceptable practice?
----- **yes .**
14. What is linear regression in statistics?
----- **Linear regression is the most basic and commonly used predictive analysis. Regression estimates are used to describe data and to explain the relationship.**
15. What are the various branches of statistics?
----- **Two branches, descriptive statistics and inferential statistics, comprise the field of statistics.**