

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

Answer: 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?

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

Answer: 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: 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

Answer: All of the mentioned

5. _____ random variables are used to model rates.

- a) Empirical
- b) Binomial
- c) Poisson
- d) All of the mentioned

Answer: Poisson

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

Answer: False

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

Answer: Hypothesis

8. 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: 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: 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?

Answer: A normal distribution is the proper term for a probability bell curve and normal distribution the mean is zero and the standard deviation is 1 and Normal distributions are symmetrical, but not all symmetrical distributions are normal

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

Answer: Data scientists use two data imputation techniques to handle missing data that is Average imputation and common-point imputation Average imputation uses the average value of the responses from other data entries to fill out missing values will recommended Average imputation

12. What is A/B testing?

Answer: It's also known as split testing or bucket testing and it is a method of comparing two versions of a webpage or app against each other to determine which one performs better

13. Is mean imputation of missing data acceptable practice?

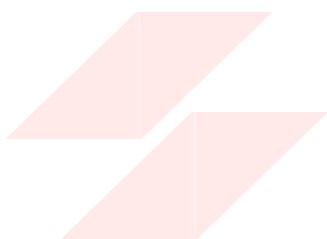
Answer: Yes imputing the mean preserves the mean of the observed data So if the data missing completely at random the estimate of the mean remains unbiased

14. What is linear regression in statistics?

Answer: linear regression is a linear approach for modelling the relationship between a scalar response and one or more explanatory variables also known as dependent and independent variables

15. What are the various branches of statistics?

Answer: I know only Psychometrics and Physics Statistics and Population Statistics



FLIP ROBO

