FLIP ROBO

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: 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?
    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

Answer: C) The square of a standard normal random variable follows what is called chi-squareddistribution

5. _____random variables are used to model rates.
    a) Empirical
    b) Binomial
    c) Poisson
    d) All of the mentioned

Answer:C) Poisson

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

Answer: b)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: a) 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: c) Outliers cannot conform to the regression relationship

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

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

Answer: Normal distributions are symmetrical, but not all symmetrical distributions are normal. A normal distribution is the proper term for a probability bell curve.

11. How do you handle missing data? What imputation techniques do you recommend?
Answer:
1. **when low data is missing, we use Mean or Median Imputation.**
2. **When large data is missing, we use Deleting Rows with missing values**

12. What is A/B testing?

Answer: 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. Eg: one company has one product and want to increase the sale so by using A/B testing . we categories the same product into two part. A as original and some changes on B. customer response on A and B . we can find the best.

13. Is mean imputation of missing data acceptable practice?

Answer: The process of replacing null values in a data collection with the data's mean is known as mean imputation. Mean imputation is typically considered bad practice since it ignores feature correlation

14. What is linear regression in statistics?
Answer: If only one independent variable in model is called simple linner model and if more than one independent variable in model is called multiple linear regression.

15. What are the various branches of statistics?
Answer:
Inferential statistics- eg-Populations
Descriptive statistics –eg- sample of population