

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

It describes events having exactly two outcomes.

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

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-

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-D

5. _____ random variables are used to model rates.
 - a) Empirical
 - b) Binomial
 - c) Poisson
 - d) All of the mentioned
6. Usually replacing the standard error by its estimated value does change the CLT.
 - a) True
 - b) False

Answer- B

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- B

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- A

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

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?
- 11. How do you handle missing data? What imputation techniques do you recommend?
- 12. What is A/B testing?
- 13. Is mean imputation of missing data acceptable practice?
- 14. What is linear regression in statistics?
- 15. What are the various branches of statistics?

10. Normal distribution is the proper term for a probability bell curve. In normal distribution the mean is zero and the SD is either Right skewed, Left skewed or Symmetric.

11. There are three techniques used to fill NaN values. These are-

- Simple
- Knn
- Iterative

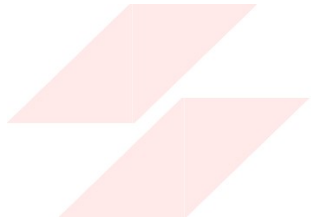
13. It is not a good practice.

14. Linear regression predicts the dependent variable using regression line based on independent variable.

$$y = a + b * x + e$$

15. The various branches of statistics are-

- Data Collection
 - Descriptive Statistics
 - Inferential Statistics
-



FLIP ROBO