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
ANS- 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
ANS- 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
ANS- 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
ANS- d) All of the mentioned

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

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

ANS-Normal Distribution, also known as the Gaussian distribution,is a probability distribution that is symmetric about the mean,showing that data near the mean are more frequent in occurrence than data far from the mean.It appears as bell curve.

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

ANS-Missing data For categorical values replace it with mode/frequent

For continuous values replace it with mean value

Otherwise drop that rows or columns.

12. What is A/B testing?

ANS-A/B testing is a type of experiment in which you split your web traffic or user base into two groups,and show two different versions of a web page,app,email,and so on,with the goal of comparing the results to find the more successful version.

13. Is mean imputation of missing data acceptable practice?

ANS-No,it is not a acceptable practice.

14. What is linear regression in statistics?

ANS- linear regression in statistics establishes the linear relationship between two variable based on line of best fit.

15. What are the various branches of statistics?

ANS-1)Descriptive Statistics

2)Inferential Statistics