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 Theorm

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 : a) The exponent of a normally distributed random variable follows what is called the log-normal distribution.

5. _____ random variables are used to model rates.

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

ANSWER : c) Poisson

6. Usually replacing the standard error by its estimated value does change the CLT.

a) True b) False

ANSWER : b) 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 : b) 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 : 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 conforms 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 : In Normal Distribution is the data is well spred in symmetric form around mean, median and mode. It is also called as bell curvrd distribution. The centre line indicate mean which is equal to 0. And standard daviation is +- 1. the data which lies after -+3 standard daviations are called outliers.

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

ANSWER : Whenever you find missing data, you can fill it by using Mean, Median or Mode techniques.

12. What is A/B testing?

ANSWER : A/B testing is a technique which is used to compare the various versions of the variables to know which version is best and effective to perform.

13. Is mean imputation of missing data acceptable practice?

ANSWER : NO, it is not acceptable during practice. it is always better to fill the empty data by calculating the mean. Because it will give the accurate results and less errors will occure during practice.

14. What is linear regression in statistics?

ANSWER : Linear regression is used to predict the value of a variable by relating some other variables of the given data. Linear regression equation is y=a+bx+e

Where, y=target variable, a=intercept, b= coefficient value, x= independent variable, e=error.

15. What are the various branches of statistics

ANSWER : STATISTICS : 1.DESCRIPTIVE STATISTICS (organising & summerising of data)

2. INFERENTIAL STATISTICS (by using data you have to measure output)

DESCRIPTIVE STATISTICS : 1. CENTRAL TENDENCY (Mean, Median, Mode)

2. DISPERSION OF DATA (Range,Variance, standard daviation, percentile)

INFERENTIAL STATISTICS : Z Score, Hypothesis testing, Z- TEST, T-TEST, Chi-Square TEST, Annova TEST