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

   e) 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

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

   Ans: 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

   e) Ans: 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

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 is one of the types of symmetrical distribution, which forms a bell curve in the plot from the set of given data. It is one of the significant techniques in statistical analyses which works under 3 key features which are mean, median and mode.

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

Ans:
There are two ways to handle the missing data one is to ignore the missing data and the other we could use and imputation technique by estimating a value based on the number of data sets which has been lost.
The best imputation technique that I would recommend is Complete Case Analysis as the method is easy to implement.

12. What is A/B testing?

Ans:
A/B testing is the process of comparing 2 versions of data and to confirm which version performs better in the provided environment to get the better efficiency.

13. Is mean imputation of missing data acceptable practice?

Ans:
No, mean imputation of missing data is not an acceptable practice

14. What is linear regression in statistics?

Ans:
Linear Regression is based on two factors i.e. dependent variable and independent variable. It is used to predict the values of the given data set, where the predicted value are the dependent variable and the values used to predict the dependent variable are called as independent variable.

15. What are the various branches of statistics?

Ans:
There are three branches in Statistics, they are Data Collection, Descriptive Statistics and inferential Statistics.