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)

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

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

Answer(B)
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)

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)

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)

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

Answer : PDF (Probability Density function) which decribes the observation different valve in distribution

There are Symmetry, Central tendency, Standard deviation, Empirical rule:
 Mention and distribution Bell curve

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

Answer
1 Complete Case Analysis
2 Mean/Mode/Median Imputation
3 Last Observation Carried Forward
4 Multiple Imputation

12. What is A/B testing?

A/B testing, also known as split testing
The process of A/B testing involves the following steps:
Hypothesis formulation
Sample selection
Testing and data collection
Statistical analysis
Conclusion and decision-making

13. Is mean imputation of missing data acceptable practice?

Mean imputation, which involves replacing missing values with the mean of the observed values for that variable, is a simple and commonly used method for handling missing data. However, whether mean imputation is an acceptable practice depends on the context and the underlying assumptions about the missingness mechanism.

14. What is linear regression in statistics?

Linear regression is a statistical technique used to model the relationship between a dependent variable and one or more independent variables. It aims to find a linear equation that best represents the relationship between the variables.

15. What are the various branches of statistics?
Descriptive Statistics

Inferential Statistics
Statistical Genetics
Probability Theory
Data Mining and Machine Learning