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

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

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

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

5. _____ random variables are used to model rates.

a) Empirical

b) Binomial

c) Poisson

d) All of the mentioned

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

a) True

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

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

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

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?

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. In graph form, normal distribution will appear as a bell curve.

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

- Complete Case Analysis(CCA):- This is a quite straightforward method of handling the Missing Data, which directly removes the rows that have missing data i.e we consider only those rows where we have complete data i.e data is not missing. ...
- Arbitrary Value Imputation. ...
- Frequent Category Imputation
-

12. What is A/B testing?

   Ans. At its most basic, A/B testing, also known as split testing, is a way to compare   different versions of something to see which performs better. In these experiments, you define a conversion goal to measure, like clicks or completed transactions. Two variations of the same marketing asset (like a web page or email) are then shown to different users at random while measuring the difference in performance

13. Is mean imputation of missing data acceptable practice?

   Ans. Mean imputation does not preserve the relationships among variables. True, **imputing the mean preserves the mean of the observed data**. So if the data are missing completely at random, the estimate of the mean remains unbiased. That's a good thing

14. What is linear regression in statistics?

 Ans. Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

15. What are the various branches of statistics?one of the mentioned

- Descriptive Statistics. In this type of statistics, the data is summarised through the given observations. ...
- Inferential Statistics. This type of statistics is used to interpret the meaning of Descriptive statistics. ...