

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: 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: 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: Modelling 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: All of the mentioned

5. _____ random variables are used to model rates.

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

Ans: Poisson

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

- a) True
- b) False

Ans: 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: Hypothesis

8. 4. Normalized data are centred at _____ and have units equal to standard deviations of the original data.

- a) 0
- b) 5
- c) 1
- d) 10

Ans: 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: Outliers cannot conform to the regression relationship.

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?

A Normal distribution is the continuous probability distribution with a probability density function that gives you a symmetrical bell curve.

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

When dealing with missing data, they tend to use two primary methods, they are imputation or data removal. Multiple mutation is powerful at dealing with missing data in multiple variables and data types.

12. What is A/B testing?

A/B testing also called as split testing is used to compare the performance of two versions of a webpage or app.

13. Is mean imputation of missing data acceptable practice?

Mean imputation does not preserve the relationships among variables. It leads to an underestimate of Standard errors, since we are mostly into studying the relationship among variables the very core is lost in mean imputation.

14. What is linear regression in statistics?

In statistics, linear regression is a statistics model which estimates the linear relationship between a scalar response and one or more explanatory variables. It is the most used predictive analysis.

15. What are the various branches of statistics?

The two main branches of statistics are descriptive statistics and inferential statistics.