

## **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 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) 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 :- c) The square of a standard normal random variable follows what is called chi-squared distribution**

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

**Answer :- c) poisson**

6. 10. 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 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?

**Answer :-** Normal distribution, also known as the Gaussian distribution or bell curve, is a probability distribution that describes how data is distributed around a mean value. The distribution is symmetric, with the majority of data points clustering around the mean, and fewer data points at the tails of the distribution.

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

**Answer :-** Missing data can be handled using imputation techniques. One common technique is mean imputation, which replaces missing values with the mean of the non-missing values. Another technique is multiple imputation, which creates multiple imputed datasets and combines the results.

12. What is A/B testing?

**Answer :-** A/B testing is a method of comparing two versions of a product or service, to determine which one performs better. It is commonly used in web design, marketing, and other fields to optimize user experience.

13. Is mean imputation of missing data acceptable practice?

**Answer :-** Mean imputation of missing data is an acceptable practice in some cases, but it can also introduce bias and lead to inaccurate results if the data is not missing completely at random.

14. What is linear regression in statistics?

**Answer :-** Linear regression is a statistical method that is used to model the relationship between a dependent variable and one or more independent variables. It involves finding the line of best fit through a set of data points, and can be used to make predictions about future data points.

15. What are the various branches of statistics?

**Answer :-** The various branches of statistics include descriptive statistics, inferential statistics, probability theory, statistical methods, applied statistics, theoretical statistics, and computational statistics.