**STATISTICS WORKSHEET-1**

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

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: - D

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: - Normal 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.

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

Answer: - There are many techniques to handle missing data. Below are the techniques.

1. Listwise Deletion
2. Pairwise deletion
3. Mean substitution
4. Regression imputation
5. Maximum likelihood
6. Expectation-Maximization
7. Multiple imputation
8. Mode substitution
9. Median substitution

There are lots of techniques to replace missing data. I suggest imputation techniques according to data.

12. What is A/B testing?

Answer: - A/B testing is a user experience research methodology. A/B tests consist of a randomized experiment with two variants, A and B. It includes application of statistical hypothesis testing or "two-sample hypothesis testing". A/B testing is a way to compare two versions of a single variable, typically by testing a subject's response to variant A against variant B, and determining which of the two variants is more effective.

13. Is mean imputation of missing data acceptable practice?

Answer: - It is not acceptable practice every time. We need to check data first.

14. What is linear regression in statistics?

Answer: - linear regression is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables.

15. What are the various branches of statistics?

Answer: - The two main branches of statistics are descriptive statistics and inferential statistics.