

## 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?

Normal Distribution is also said as Gaussian distribution is symmetric about mean about showing data near the mean. This is a proper term for a probability bell curve.

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

Missing data can be handled using techniques called imputation. In this imputation is used to replacing the missing data with suitable values to retain the most of information of data sheet. I use CCA (Complete Case Analysis) which removes the row that have missing data.

12. What is A/B testing?

A/B testing is the popular way to test your products. Is basic randomized control experiment in which we have to compare two versions of variables to find which perform better.

13. Is mean imputation of missing data acceptable practice?

Yes. Imputation is nothing but process of replacing missing data with substituted values.

14. What is linear regression in statistics?

It is used to estimate the relationship between two variables, used to know how strong the relationship is between two variables.

15. What are the various branches of statistics?

There are three major branches of statistics data collection, descriptive statistics and inferential statistics.

 **FLIP ROBO** **FLIP ROBO**

---