

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)

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

a) True
b) False

Answer: b)

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)

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)

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)

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 is the probability distribution where all values fall symmetric to the mean value (mean is at center). It is also known as Bell curve due to its shape.

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

Answer: 1. Filling average values in missing places.

2. Filling common values.

3. Using deletion method.

4. Using regression analysis.

12. What is A/B testing?

Answer: It is comparing of two variables by keeping one constant and making changes to other one based on the customer feedback or in other words testing a sample from customers.

13. Is mean imputation of missing data acceptable practice?

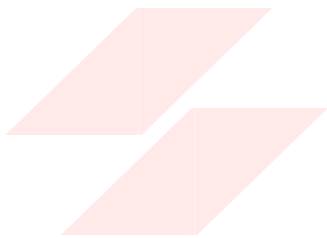
Answer: No it's an acceptable practice as it does not retain connection between variables, reduces variance.

14. What is linear regression in statistics?

Answer: It is the prediction of values of one variable based on the values of another variable.

15. What are the various branches of statistics?

Answer: There are two main branches Descriptive statistics and Inferential statistics.



FLIP ROBO