

STATISTICS WORKSHEET-1

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.-Modeling bounced 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 centered 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

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly. 10. What do you understand by the term Normal Distribution?

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

Ans.-we can use either is NA() or is Null() method to detect missing values in data.

We can get the total missing value in each column with sum() or take the average with mean().

12. What is A/B testing?

Ans.-A/B testing is the process of comparing two variations of a page element, usually by testing users

Response to variant A vs. Variant B and concluding which of the two variant is more effective.

13. Is mean imputation of missing data acceptable practice?

Ans.-Mean imputation is typically considered terrible practice since it ignores feature correlation.

14. What is linear regression in statistics?

Ans.-Linear regression analysis is used to predict the value of a variable based on the value of another

Variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called independent variable.

15. What are the various branches of statistics

Ans.- There are three branches of statistics

1.- Data collection

2.-Descriptive Statistics

3.-inferential Statistics