

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: c

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

10. What do you understand by the term Normal Distribution?

Answer: Normal distribution is a fundamental concept in statistics and probability theory . It describes the probability distribution of a continuous random variable where the data clusters symmetrically

around the mean. It is widely used in various fields like finance, science etc because many concept in this used to follow the same fundamental.

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

Answer: Handling missing data is a crucial step in data analysis because incomplete data can lead to biased results and inaccurate conclusions. There are several techniques for handling missing data, the techniques are: forward/backward fill, mean/median/mode imputation, k-nearest neighbors imputation, multiple imputation, linear interpolation, predictive modeling, hot desk imputation. The choice of imputation techniques will depend on the specific characteristics of the dataset.

12. What is A/B testing?

Answer: A/B testing is a method use to compare two version of product to determine which is best. It is commonly used in marketing, development and website optimization. It allows business to make improvements to their products.

13. Is mean imputation of missing data acceptable practice?

14. What is linear regression in statistics?

Answer: It is a statistical method used to model the relationship between one or more independent variables and a dependent variable . It assumes a linear relationship between the independent variables and the dependent variable, which can be represented by a straight line in a scatter plot.

15. What are the various branches of statistics?