Statistics worksheet1

1.Bernoulli random variables take (only) the values 1 and 0

Ans: a) 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?

Ans :d)all of the mentioned

3. Which of the following is incorrect with respect to use of Poisson distribution?

Ans: modelling bounded 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

5. _____ random variables are used to model rates.

Ans: c) Poisson

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

Ans : a) True

7. 1. Which of the following testing is concerned with making decisions using data?

Ans: b) Hypothesis

8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original data.

Ans: c) 1

9. Which of the following statement is incorrect with respect to outliers?

Ans: c) Outliers cannot conform to the regression relationship

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

Normal distribution, also known as the Gaussian 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?

Use deletion method to eliminate missing data. The deletion methods only work for certain datasets where participants having missing fields

Use regression analysis to systematically eliminate data

Imputation techniques:

1.Complete case analysis (CCA)

2.Arbitrary value Imputation

3.Frequent category Imputation

12. What is A/B testing?

A/B testing is a method of comparing two versions of a webpage or app against each other to determine which one performs better.

13. Is mean imputation of missing data acceptable practice?

Mean imputation does not preserve the relationships among variables. If the data are missing completely at random, the estimate of the mean remains unbiased.

14. What is linear regression in statistics?

Linear regression is a linear approach for modelling the relationship between the scalar response and one or more explanatory variables(ie., dependent and independent variables).

15. What are the various branches of statistics?

There are 3 real branches of statistics.

1. Data collection
2. Descriptive statistics
3. Inferential statistics