

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) FalseAns: (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?
 - a) Central Limit Theorem
 - b) Central Mean Theorem
 - c) Centroid Limit Theorem
 - d) All of the mentionedAns: (a) 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 mentionedAns: (b) Modeling 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 mentionedAns: (d) All of the mentioned
5. _____ random variables are used to model rates.
 - a) Empirical
 - b) Binomial
 - c) Poisson
 - d) All of the mentionede) Ans: (c) Poisson
6. 10. Usually replacing the standard error by its estimated value does change the CLT.
 - a) True
 - b) FalseAns: (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 mentionedAns:(b) 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: (a) 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: (c) Outliers cannot conform to the regression relationship

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?

Ans: Normal distribution is nothing but mean and variance distribution

Normal distribution mean=median=mode

Symmetry about the center

The curve of distribution is bell shaped and symmetrical about the line and total area under the curve is 1.

It represents behavior most of the situation

In this half values is left to the center and half is right to the center

In this not required whole data for plotting or distribution just required mean and standard deviation

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

Ans: Missing values means miss some data in to the dataset data imputation technique is one of that that technique to handle the missing value or data into the dataset. Imputation technique filling missing value based on other data. One option involves replacing any missing value with the mean or median of that variable for all other observations, which has the benefit of not changing the sample mean for that variable.

In that there are two techniques: Generalized imputation, similar case imputation.

For example: In this technique we can impute the missing value by calculating the mean of the available Values and after calculating whatever value x is calculate as a mean that x is consider as a missing value. Similar case technique calculate the mean by individually calculating mean by other column similar and Whatever value we will wet calculate impute this value in to missing data place .

12. What is A/B testing?

Ans: A/B testing is used to compare two different products and it is depend on user input

A/B testing is most of the companies to selling their products. When we applying some machine

learning algorithm in E commerce site for example Amazon I am buying a phone and suddenly

I am getting recommendation of ear phone, headphones and different types of products so what happen

in that we may attracted towards that and we go and click it so basically happens that next time I am

visit that particular site they have that particular information that that person like that product because

he clicks on that and order or not that information is also stored into that. So A/B testing is heavily use

in E commerce site different product dynamically popup in front of you. Actually testing in data science

they usually have that content A/B testing what type content is to be display user is actually click in it

that clicks is basically recorded is saved in somewhere databases and all they particular information is

track using this information many companies use that to increase sell there products

13. Is mean imputation of missing data acceptable practice?

Ans: Mean imputation does not preserve the relationships among variables. So

the data are missing completely at random, the estimate of the mean remains unbiased. That's a good thing.

Plus, by imputing the mean, you are able to keep your sample size up to the full sample size

14. What is linear regression in statistics?

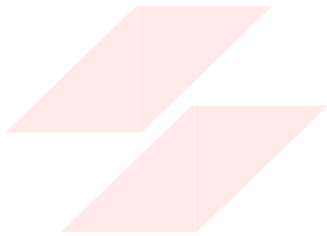
Ans: It is statistical way measuring the relationship between variable In statistics, linear regression is a linear approach to modeling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables). The case of one explanatory variable is called simple linear regression; for more than one, the process is called multiple linear regression. This term is distinct from multivariate linear regression, where multiple correlated dependent variables are predicted, rather than single scalar variable.

15. What are the various branches of statistics?

Ans: There are two branches of statistics Descriptive and Inferential

Descriptive: Organizing, Summarizing, Presenting data in a informative way

Inferential: It is used to determining the something about population on basis of sample.



FLIP ROBO