Statistics worksheet 1

Ans 1- A,(true)

Ans 2- A,(Central limit theorem

Ans 3- B,(Modelling bounded count data)

Ans 4- D,(all of the mentioned)

Ans 5- C,(Poisson)

Ans 6- B,(False)

Ans 7- B,(Hypothesis)

Ans 8- A,(0)

Ans 9- C,(Outliers cannot conform to the regression relationship)

Ans 10- Normal distribution is also known as gaussian distribution. In statistics it is the continuous probality distribution it can be defined as probability density function for a continuous random variable. It forms a bell curve when plotted and it can be used in a real world ex- like in the competitive exams here, the normal distribution will define the max students will score the average marks, while smaller amount of students will score the grades (b,d) and a smaller amt of students score will be (a,f) and it can be derived from the empirical rule or formula of the normal distribution curve. Normal distribution posses some very ossom properties like It is symmetrical, this means we can divide the curve in 2 equal parts.

If a data is skewed data it can be made normal by presenting it inside the normal distribution's bell curve as normal distribution have no skewness in it.

Ans 11-

Ans 12- It is a hypothesis testing for a randomized experiment with two variables A and B.

The goal of A/B Testing is to identify any changes to the web page to maximize or increase the outcome of interest. A/B testing is a fantastic method for figuring out the best online promotional and marketing strategies for your business. It can be used to test everything from website copy to sales emails to search adsAn example of this could be identifying the click-through rate for a banner ad.

Ans 13- If we talk about filling the Null values then, we can easily ans that we have many methods to do the same, from their we have one technique where we fill the, NA values with the help of, calculating their mean values. so now the questions If we are using this technique to handle null values is it a good practice or not. As we know all the techniques has some more or less pros and cons so in this case we have many pros thus we can surely use this method. if we are using this technique we will safe in case of data-loss from the dataset which might be possible in dropping function but the only cons is that it not useful for all type of datasets we have choose this replacing technique according to type of dataset otherwise, this may further can creating a problem in making ML models.

Ans 14- A linear regression is a good tool for quick predictive analysis: for example, the price of a house depends on a myriad of factors, such as its size or its location. In order to see the relationship between these variables, we need to build a linear regression, which predicts the line of best fit between them and can help conclude whether or not these two factors have a positive or negative relationship.

Ans 15- In statistics we have two branches in total which gives us the definition of how stats work and how important it is for us statistics is a part of mathematics used to perform different operations - Data collection, analysis, and so on. Statistics examine the

methodology of collecting, reviewing, and making Data conclusions. Thus the two branches of statistics are:

a.) Descriptive statistics: In this type of statistics data is summarized through the given observations. The summarization is in form of sample of population using parameters such as mean or standard deviation. descriptive statistics is the way to collect, organize and display the data using the table, graphs and summary measures.

b.)Inferential statistics:- The type of statistics is used to interpret the conclusions of the Descriptive stats. That means as the descriptive stats is used to collect, organize and display data on the other hand inferential stats is used to, analyze that data and get some insights.It allows us to use information collected from the samples to make decisions, predictions or a inferences from a population.