Q.1 a) True

Q.2 a) Central Limit Theorem

Q.3 c) Modelling contingency tables

Q.4 d) All of the mentioned

Q.5 c) Poision

Q.6 b) False

Q.7 b) Hypothesis

Q.8 a) 0

Q.9 c) Outliers cannot confirn to the regression relationship

Q.10 Normal distribution is also called Gaussian distribution. It is most commonly seen continuous distribution in nature. In Normal distribution the mean, median and mode all line up such that the center of distribution is the mean.

Q.11 Handling missing data is important aspect od data analysis and modelling. We can delete the entire row if it contains all missing values.

Mean/Median imputation : Replacing missing value with mean or median of the values of that variable.

Mode imputation : For categorical data, we can replace the missing values with mode of that variables.

Q.12    A/B testing is a method of comparing two versions of app feature, webpage to determine which version lead to better result. It is commonly used in marketing, product development,etc. It is also known as split testing.

Q.13    Mean imputation is used to handle the missing data especially when the data is minimal or random. It helps in maintain basic statistics of the dataset. Mean imputation can introduce bias in dataset if the missing data is not missing completely at random. It is generally not recommended for datasets when the missing data is not random.

Q.14    Linear regression is a statistical method used to generate a relationship between dependent variable and one or more independent variables. The goal is to find the best fitting linear relationship.

Mathematically,

$Y = b_o + b_1 x_1 + b_2 x_2 + \ldots\ldots + b_n x_n$

Where y is the dependent variable

Bo is the intercept

B1,b2….. bn are coefficients of the independent variables

X1,x2, …. Xn are the independent variables

Q.15　　Various branches of statistics are:

1.Descriptive Statistics : It provides an overview of the datas main characteristics using mean, median, mode, standard deviation and histograms.

2. Inferential Statistics : It involves conclusion about population based samples.

3. Probability Theory : It forms the foundation of statistical methods and provide mathematical framework to analyze random events.

4. Biostatistics : It is used for biological and medical data.