ASSUMPTIONS OF LINEAR REGRESSION

Linear regression is a statistical technique used to model the relationship between a dependent variable (usually denoted by "y") and one or more independent variables (usually denoted by"x").

The assumptions are:

1.Linearity:

It states that the dependent variable y should be linearly related to independent variable x.This assumption can be checked by plotting a scatterplot between both variables.

2.Normality:

The X and Y variable should be normally distributed.Histograms,KDE plots,Q-Q plots can be used to check the normality assumption.

3.Homoscedasticity:

The variance of the error terms should be constant.i.e:The spread of residuals should be constant for all values of X.This assumption can be checked by plotting a residual plot.If the assumption is violated then the points will form a funnel shape otherwise they are constant.

4.no multicollinearity:

The variables should be independent of each other. I.e no correlation should be there between the independent variable.To check the assumption,we can use a correlation matrix or VIF score. If the VIF score is greater than 5 then variables are highly correlated

5.Error terms should be normally distributed .

q-q plots and histograms can be used to check distribution of error terms.

6. No Autocorrelation:

The error terms should be independent of each other. Auto correlation can be tested using Durbin watson test. The null hypothesis assumes there is no auto correlation. The value of the test lies between 0 to 4. If the value of the test is 2 ,then there is no autocorrelation.