

Multicollinearity

What do you mean by the problem of multicollinearity? What are the causes of multicollinearity?

Ans:- Multicollinearity is used to denote the presence of linear relationship between explanatory variables in a regression model. In the OLS method one assumption is that explanatory variables should not be perfectly correlated to other explanatory variables. The problem of multicollinearity arises whenever, either in the population or in the sample, various explanatory variables are in exact linear relationship with each other. On the other hand when all explanatory variables are uncorrelated (or independent of one another) with each other, then there is absence of multicollinearity. The cases between this two extremes are referred to as the ^{various} degree of multicollinearity.

* Multicollinearity indicates a strong linear relationship among the explanatory variables in a multiple regression model.

CAUSES OF M.C.

There are various reasons for causing the problem of multicollinearity. They are -

1) There is a tendency of economic variables to move together over time. For example, in time

series data, growth and trend factors are the main causes of multicollinearity. In period of rapid economic growth and boom, the income, consumption, savings, investment, prices, employment tend to rise and decrease in periods of recession.

2) Multicollinearity arises due to use of lagged values of certain explanatory variables in the regression model as observed in distributed lagged model. For example, in consumption function, lagged income (Y_{t-1}) is also taken as one of the independent variable such that $C_t = f(Y_t, Y_{t-1})$. The successive values of a certain variables are interrelated i.e., income in current period is partly determined by its own value in the previous period, and so on. Thus multicollinearity exists in distributed lagged model.

3) Generally multicollinearity is associated with time series analysis. It arises in a cross section studies also. If a sample of household is such that high income and high capital assets got together, the separate influences of income and assets on consumption can not be estimated since both of them are correlated. Multicollinearity tends to be more common and serious problem in time series.