Coefficient of Determination R^2

In statistics, the **coefficient of determination**, R^2, is used in case of statistical models whose main purpose is the prediction of future outcomes on the basis of other related information. It is the proportion of variability in a data set that is accounted for by the statistical model It provides a measure of how well future outcomes are likely to be predicted by the model.

R^2 is a statistic that will give some information about the goodness of fit of a model. In regression, the R^2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An R^2 of 1.0 indicates that the regression line perfectly fits the data.

R^2 is often interpreted as the proportion of response variation "explained" by the regressors in the model. Thus, $R^2 = 1$ indicates that the fitted model explains all variability in y, while $R^2 = 0$ indicates no 'linear' relationship between the dependent variable and independent variables. A value such as $R^2 = 0.7$ may be interpreted as follows: "Approximately seventy percent of the variation in the dependent variable can be explained by the independent variable. The remaining thirty percent can be explained by unknown, variables ."

A caution that applies to R^2, as to other statistical descriptions of correlation and association is that "correlation does not imply causation." In other words, while correlations may provide valuable clues regarding causal relationships among variables, a high correlation between two variables does not represent adequate evidence that changing one variable has resulted, or may result, from changes of other variables.

In case of a single independent variable, fitted by least squares, R^2 is the square of the Pearson product-moment correlation coefficient relating the independent and the dependent variable. More generally, R^2 is the square of the correlation between the independent and the dependent variables.
