**linear Regression**

Linear regression (statistics) is a linear approach for modelling the relationship between a dependent variable and independent variables (explanatory variables). If we have only one independent variable, the approach is called simple linear regression. Simple linear regression models the relationship between two variables by fitting a linear equation to observed data. For example, if we want to relate the weights of individuals to their heights using a linear regression model. In this simple case a linear regression line has an equation: Y = a + bX, where X is the explanatory variable and Y is the dependent variable. The slope of the line is b, and a is the intercept (the value of y at x = 0). In case of more than one independent variables, the process is called multiple linear regression.

Before attempting to fit a linear model to observed data, a we should first determine whether there is a relationship between the variables of interest. This does not necessarily imply that one variable causes the other, but that there can be some significant association between the two variables. If there is no association between the explanatory and dependent variables, then fitting a linear regression model probably will not be useful. A valuable numerical measure of association between two variables is the correlation coefficient. Its value can be in range between -1 and 1, which indicates the strength of the association of the observed data for the two variables.

The relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data (linear models). Linear regression focuses on the conditional probability distribution of the response given the values of the predictors.

Practical application of linear regression:

1. Prediction, or error reduction. Linear regression is used to fit a predictive model to an observed data set of values of the response (dependent variable) and explanatory variables (independent variables).
2. Explaining variation in the response variable that can be attributed to variation in the explanatory variables. In this case linear regression is applied to quantify the strength of the relationship between the response and the explanatory variables; to determine whether some explanatory variables may have no linear relationship with the response.

Linear regression models are often fitted using the least squares approach. However, they hey may also be fitted as by minimizing the "lack of fit" in some other norm or by minimizing a penalized version of the least squares cost function as in ridge regression (L2-norm penalty) and lasso (L1-norm penalty). The least squares approach can be used to fit models that are not linear models.

**Sources:**

1. <https://en.wikipedia.org/wiki/Linear_regression>
2. http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm