Assignment #3: Linear regression/Linear fit

Biol6930

Ling Wang

September 13, 2018

Statistically, linear regression is a predictive analysis through finding the best fit line to forecast future values. It is a linear approach to modeling the relationship between a dependent variable and on or more independent variables. One independent variable involved linear regression is called simple linear regression. Simple linear regression allows us to summarize and study relationships between two continuous variables. For more than one independent variables, the case is called multiple linear regression. Estimations of those regressions contribute to explain the relationship between on dependent variable and one or more independent variables, such as does a set of predictor variables predicting a dependent variable well; and which variables are significant predictors of the dependent variables. The simplest form of the regression equation with one dependent and one independent variable is defined by the formula y = c + b\*x, where y = estimated dependent variable score, c = constant, b = regression coefficient, and x = score on the independent variable.

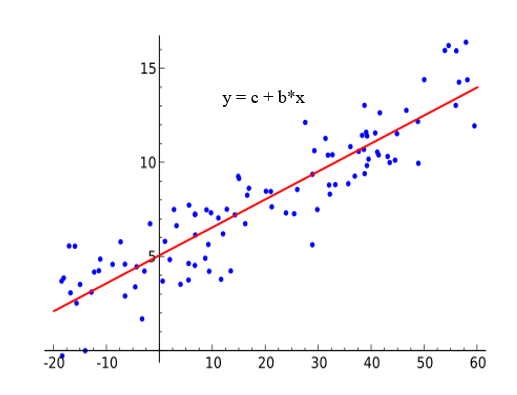


Figure 1: Regression Analysis Model

The best fit line on a scatter graph is found and used as the basis for estimating the future estimating values of the function. While, the model for multiple linear regression, given n observations, is y=b0+b1xi1+b2xi2+…bnxin+c, for i=1, 2, …n.

The relationships in linear regression are modeled by linear predictor functions with unknown model parameters are estimated from the data. The regression might be used to identify how strong of an effect that the independent variable(s) impacts on a dependent variable.  Also, it can be used to helps us to understand how much the dependent variable changes with a change in one or more independent variables.  Additionally, regression analysis predicts trends and future values by extending the best fit line while maintaining its slope.