

Regression Analysis

Regression analysis uses mathematical curves to summarize relationships among several variables. It is somewhat similar to correlation analysis. But correlation analysis has two major shortcomings.

Correlation coefficients reveal only linearity in relationships; and correlation coefficients do not yield models that reveal precisely how one variable affects another.

Regression analysis works by creating a line of best fit by minimizing the sum of squares of predictive errors. The resulting regression line can be considered a running mean (a line that estimates the mean of the dependent variable from for particular values of the independent variables).

Regression analysis results in equations that that give precise formulas that tell how and to what degree a dependent variable is affected by independent variables. However, it should be understood that the dependent variable (usually referred to as "Y") is but an estimated mean value, and is therefore subject to error.

Regression analysis is instrumental in uncovering and modeling complex relationships where one variable is a function of multiple variables. For instance, if an investor felt that the price-to-book-value was a function of equity and return-on-equity ($\frac{Price}{Book\ Equity} = Book\ Equity * ROE$), that investor would perform regression analysis with P/B designated as the dependent variable, and Book Equity and ROE designated as the independent variables.