

Multiple regression essay



**ASSIGN
BUSTER**

Multiple regression, a time-honored technique going back to Pearson's 1908 use of it, is employed to account for (predict) the variance in an interval dependent, based on linear combinations of interval, dichotomous, or dummy independent variables. Multiple regression can establish that a set of independent variables explains a proportion of the variance in a dependent variable at a significant level (through a significance test of R^2), and can establish the relative predictive importance of the independent variables (by comparing beta weights). Power terms can be added as independent variables to explore curvilinear effects.

Cross-product terms can be added as independent variables to explore interaction effects. One can test the significance of difference of two R^2 's to determine if adding an independent variable to the model helps significantly. Using hierarchical regression, one can see how most variance in the dependent can be explained by one or a set of new independent variables, over and above that explained by an earlier set. Of course, the estimates (b coefficients and constant) can be used to construct a prediction equation and generate predicted scores on a variable for further analysis.

The multiple regression equation takes the form $y = b_1x_1 + b_2x_2 + \dots + b_nx_n + c$. The b's are the regression coefficients, representing the amount the dependent variable y changes when the corresponding independent changes 1 unit.

The c is the constant, where the regression line intercepts the y axis, representing the amount the dependent y will be when all the independent variables are 0. The standardized version of the b coefficients are the beta

weights, and the ratio of the beta coefficients is the ratio of the relative predictive power of the independent variables. Associated with multiple regression is R^2 , multiple correlation, which is the percent of variance in the dependent variable explained collectively