Partial and Semi-Partial Correlations

Partial Correlation

One is often interested in knowing what the product-moment correlation would be between two variables if one or more related variables could be held constant. For example, in one of our previous examples, we may be curious to know what the correlation between achievement in learning French is with past achievement in learning English with intelligence held constant. In other words, if that proportion of variance shared by both French and English learning with IQ is removed, what is the remaining variance shared by English and French?

When one subtracts the contribution of a variable, say, X_3, from both variables of a correlation say, X_1 and X_2, we call the result the partial correlation of X_1 with X_2 partialling out X_3. Symbolically this is written as $r_{12.3}$ and may be computed by

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

More than one variable may be partialled from two variables. For example, we may wish to know the correlation between English and French achievement partialling both IQ and previous Grade Point Average. A general formula for multiple partial correlation is given by

$$r_{12.34..k} = \frac{(1.0 - R^2_{y.34\,..\,k}) - (1.0 - R^2_{y.12\,..\,k})}{1.0 - R^2_{y.34\,..\,k}}$$

Semi-Partial Correlation

It is not necessary to partial out the variance of a third variable from both variables of a correlation. It may be the interest of the researcher to partial a third variable from only one of the other variables. For example, a researcher may be interested in the correlation between grade point average and a standardized achievement test after the intelligence score of the subjects is removed from one or both variables. When the variance of a third variable is partialled from only one of the variables in a correlation, we call the result a semi_partial or part correlation. The symbol and calculation of the part correlation is

$$r_{1(2.3)} = \frac{r_{1,2} - r_{1,3}r_{2,3}}{\sqrt{(1.0 - r^2_{23})}}$$

where X_3 is partialled only from X_2 .

The squared multiple correlation coefficient R^2 may also be expressed in terms of semi_partial correlations. For example, we may write the equation

$R^2_{y.1\ 2\ ..\ k} = r^2_{y.1} + r^2_{y(2.1)} + r^2_{y(3.12)} + .. + r^2_{y(k.12..k-1)}$

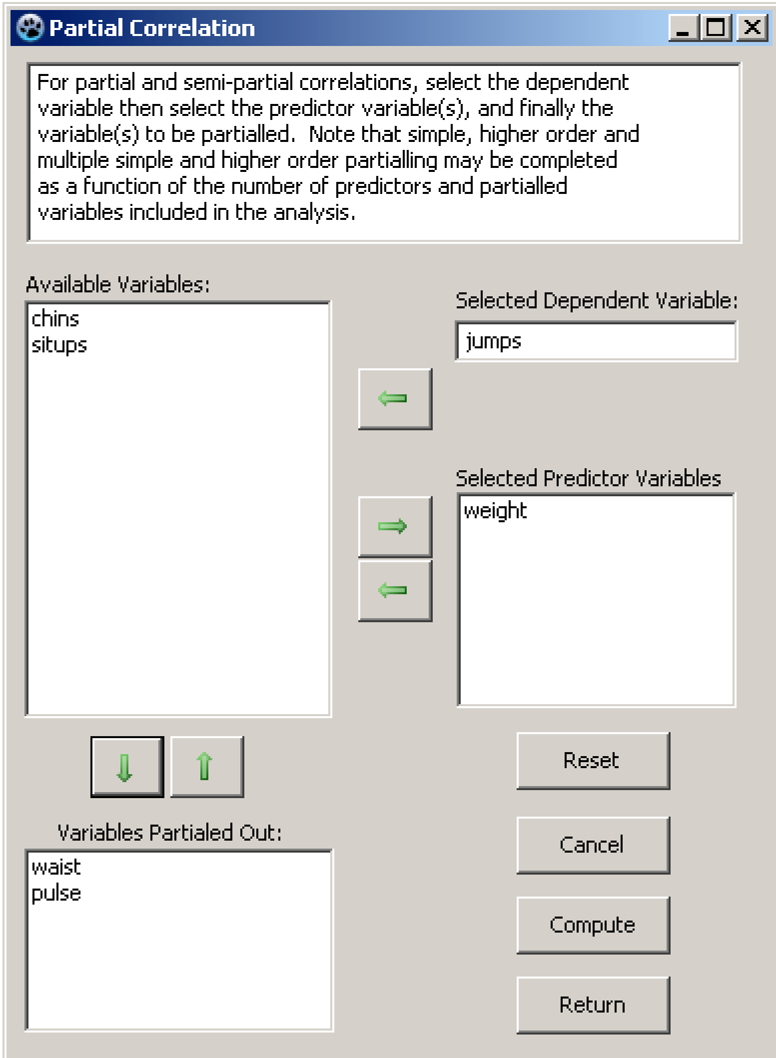
In this formula, each semi_partial correlation reflects the proportion of variance contributed by a variable independent of previous variables already entered in the equation. However, the order of entry is important. Any given variable may explain a different proportion of variance of the independent variable when entered first, say, rather than last!

The semi-partial correlation of two variables in which the effects of K-1 other variables have been partialed from the second variable may be obtained by multiple regression. That is

$r^2_{y(1.2\ 3\ ..\ k)} = R_{2y.1\ 2\ ..\ k} - R^2_{y.23..k}$

To demonstrate partial and semi-partial correlations, we will use the cansas.LAZ file. We will want to examine the relationship between weight and jumps when the influence of waist and / or pulse is removed.

The dialog is shown below:



Partial and Semi-Partial Correlation Analysis

Dependent variable = jumps

Predictor Variables:
Variable 2 = weight

Control Variables:
Variable 2 = waist
Variable 3 = pulse

Multiple partialling with 2 variables.

Squared Multiple Correlation with all variables = 0.054

Standardized Regression Coefficients:
weight = -0.259
waist = 0.015
pulse = -0.055

Squared Multiple Correlation with control variables = 0.038

Standardized Regression Coefficients:
waist = -0.205
pulse = -0.037

Partial Correlation = 0.129

Semi-Partial Correlation = 0.127

F = 0.271 with probability = 0.6099, D.F.1 = 1 and D.F.2 = 16