有關 Correlations 的補充資料

1. PARTIAL CORRELATION

Def: A partial correlation is a correlation between two variables from which the linear relations, or effects, of another variable(s) have been removed. The partial correlation is symmetric, i.e., $r_{12.3} = r_{21.3}$.

Specification:

The partial correlation between two variables when one variable is partialed out is called a ***first-order partial correlation***. For example, $r_{12.34}$ is the ***second-order partial correlation*** between variables 1 and 2 from which 3 and 4 were partialed out. And $r_{12.345}$ is a ***third-order partial correlation***. The correlation between two variables from which no other variables are partialed out is called a ***zero-order correlation***, such as: $r_{12},\ r_{13},\ \textit{and}\ r_{23}$. It is possible for the sign of the partial correlation to differ from the sign of the zero-order correlation coefficient between the same variables. Also, the partial correlation coefficient may be larger or smaller than the zero-order correlation coefficient between the variables.

The squared partial correlation coefficient is a ratio of variance incremented to residual variance.

Ex:

1st-order partial corr.: $$r_{12.3} = r_{e_1e_2} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$$

2nd-order partial corr.: $$r_{12.34} = \frac{r_{12.3} - r_{14.3}r_{24.3}}{\sqrt{1-r_{14.3}^2}\sqrt{1-r_{24.3}^2}}$$

$$r_{12.3}^2 = \frac{SSE(X_3) - SSE(X_2, X_3)}{SSE(X_3)} = \frac{SSR(X_2 \mid X_3)}{SSE(X_3)}$$

$$= \frac{R_{1.23}^2 - R_{1.3}^2}{1 - R_{1.3}^2}$$

$$= \frac{R_{2.13}^2 - R_{2.3}^2}{1 - R_{2.3}^2}$$

$$r_{12.34}^2 = \frac{R_{1.234}^2 - R_{1.34}^2}{1 - R_{1.34}^2} = \frac{R_{2.134}^2 - R_{2.34}^2}{1 - R_{2.34}^2} = \frac{SSR(X_2 \mid X_3, X_4)}{SSE(X_3, X_4)}$$

$$r^2_{12.345} = \frac{R^2_{1.2345} - R^2_{1.345}}{1 - R^2_{1.345}} = \frac{R^2_{2.1345} - R^2_{2.345}}{1 - R^2_{2.345}}$$

2. SEMIPARTIAL CORRELATION

Def: A semipartial correlation is a correlation between an unmodified variable and a variable that was residualized. The symbol for a first-order semipartial correlation is $r_{1(2.3)}$, which means the correlation between X_1 (unmodified) and X_2, after it was residualized on X_3, or after X_3 was partialed out from X_2.

Specification:

$r_{12.3}$ will be larger than either $r_{1(2.3)}$ *or* $r_{2(1.3)}$, except when r_{13} and r_{23} equals zero, in which case the partial correlation will be equal to the semipartial correlation. When $r_{13} = r_{23}$, the two squared semipartial correlations yield the same resuls. When $|r_{13}| > |r_{23}|$, then $r^2_{2(1.3)} > r^2_{1(2.3)}$. The converse is , of course, true when $|r_{13}| < |r_{23}|$.

The squared semipartial correlation is also called ***part correlation***. A squared semipartial correlation indicates the proportion of variance of the dependent variable accounted for a given independent variable after another variable(s) has already been taken into account. Usually speaking, the partial correlation is larger than its corresponding semipartial correlation.

The sign of the semipartial correlation is the same as the sign of the regression coefficient (b or β) that corresponds to it.

Ex:

1st-order semipartial corr.: $$r_{1(2.3)} = r_{x_1 e_2} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{1 - r^2_{23}}}$$

$$r_{2(1.3)} = r_{x_2 e_1} = \frac{r_{12} - r_{13} r_{23}}{\sqrt{1 - r^2_{13}}}$$

$$r^2_{1(2.3)} = R^2_{1.23} - R^2_{1.3}$$

$$r^2_{1(2.34)} = R^2_{1.234} - R^2_{1.34}$$

$$r^2_{3(1.245)} = R^2_{3.1245} - R^2_{3.245}$$

$$R^2_{y.1234} = R^2_{y.1} + \left(R^2_{y.12} - R^2_{y.1}\right) + \left(R^2_{y.123} - R^2_{y.12}\right) + \left(R^2_{y.1234} - R^2_{y.123}\right)$$
$$= r^2_{y1} + r^2_{y(2.1)} + r^2_{y(3.12)} + r^2_{y(4.123)}$$

From the preceding examples it should be clear that to calculate a squared semipartial correlation of any order, it is necessary to (1) calculate the squared multiple correlation of the dependent variable with all the independent variables, (2) calculate the squared multiple correlation of the dependent variable with the variables that are being partialed out, (3) subtract the R^2 of step 2 from the R^2 of step 1.

3. TEST OF SIGNIFICANCE FOR SQUARED PARTIAL AND SQUARED SEMIPARTIAL CORRELATIONS

$$F = \frac{\left[\left(R^2_{y.12\ldots k_1}\right) - \left(R^2_{y.12\ldots k_2}\right)\right] / \left(k_1 - k_2\right)}{\left(1 - R^2_{y.12\ldots k_1}\right) / \left(N - k_1 - 1\right)}$$

where $R^2_{y.12\ldots k_1}$ = squared multiple correlation coefficient for the regression of Y on k1 variables (the larger coefficient); $R^2_{y.12\ldots k_2}$ = squared multiple correlation coefficient for the regression of Y on k_2 variables; k_2 = the smaller set of variables selected from among those of k_1; and N=sample size. The F ratio has *$(k_1 - k_2)$ df* for the numerator and *$(N - k_1 - 1)$ df* for the denominator.

4. MULTIPLE PARTIAL AND SEMIPARTIAL CORRELATIONS

Def: A ***multiple partial correlation*** may be used to calculate the squared multiple correlation of a dependent variable with a set of independent variables after controlling, or partialing out, the effects of another variable, or variables, from the dependent as well as the independent variables. $R^2_{1.23(4)}$ means the squared multiple correlation of X_1 with X_2 and X_3, after X_4 was partialed out from the other variables. Note that the variable that is partialed out is placed in parenthesis. Similarly, $R^2_{y.23(45)}$ is the squared multiple correlation of X_1 with X_2 and X_3, after X_4 and X_5 were partialed out from the other three variables.

Ex:

$$R^2_{1.23(4)} = \frac{R^2_{1.234} - R^2_{1.4}}{1 - R^2_{1.4}}$$

$$R^2_{1.23(45)} = \frac{R^2_{1.2345} - R^2_{1.45}}{1 - R^2_{1.45}}$$

To calculate a squared multiple partial correlation, then, (1) calculate the squared multiple correlation of the dependent variable with the remaining variables (i.e., the independent and the control variables); (2) calculate the squared multiple correlation of the dependent variable with the control variables only; (3) subtract the R^2 obtained in step 2 from the R^2 obtained in step 1; and (4) divide the value obtained in step 3 by one minus the R^2 obtained in step 2.

Def: A ***squared multiple semipartial correlation*** may be calculated from the squared multiple correlation of a dependent variable with a set of independent variables after controlling, or partialing out, the effects of another variable or variables, from the independent variables only. The notation is $R^2_{1(23.4)}$. The dependent variable is outside the parenthesis. The control variable (or variables) is placed after the dot. Similarly, $R^2_{1(23.45)}$ is the squared multiple semipartial correlation of X_1 with X_2 and X_3, after X_4 and X_5 were partialed out from X_2 and X_3.

Ex:

$$R^2_{1(23.4)} = R^2_{1.234} - R^2_{1.4}$$

$$R^2_{1(23.45)} = R^2_{1.2345} - R^2_{1.45}$$

where $R^2_{1(23.4)}$ indictes the proportion of variance in X1 accounted for by X2 and X3, after the contribution of X4 was taken into account.

Test of significance for squared multiple partoial and squared multiple semipartial correlations:

Ex: The test of $R^2_{1(23.4)}$ is as follows, it is also a test of $R^2_{1.23(4)}$. Same to the case, as the test of $R^2_{1(23.45)}$ is also use to test of $R^2_{1.23(45)}$.

$$F = \frac{\left[\left(R^2_{y.12...k_1}\right) - \left(R^2_{y.12...k_2}\right)\right] / \left(k_1 - k_2\right)}{\left(1 - R^2_{y.12...k_1}\right) / \left(N - k_1 - 1\right)}$$

where $R^2_{y.12...k_1}$ = squared multiple correlation coefficient for the regression of Y on k1 variables (the larger coefficient); $R^2_{y.12...k_2}$ = squared multiple correlation coefficient for the regression of Y on k_2 variables; k_2 = the smaller set of variables selected from among those of k_1; and N=sample size. The F ratio has *(k_1 –k_2) df* for the numerator and *(N –k_1 -1) df* for the denominator.

Reference

Pedhazur, E. J. (1997). ***Multiple regression in behavioral research: Explanation and prediction*** (pp. 160-193) (3rd ed.). Fort Worth, FL: Harcourt Brace College Publishers.