PARTIAL CORR

Notation

The following notation is used throughout this chapter unless otherwise stated:

N	Number of cases
X_{kl}	Value of variable k for case l
w_l	Weight for case l
W_{ij}	Sum of the weights of cases used in computation of statistics for variable i and j
W_i	Sum of the weights of cases used in computation of statistics for variable i

Statistics

Zero-Order Correlations

$$r_{ij} = \frac{\sum_{l=1}^{N} w_l X_{il} X_{jl} - \left(\sum_{l=1}^{N} w_l X_{il}\right)\left(\sum_{l=1}^{N} w_l X_{jl}\right) \Big/ W_{ij}}{\sqrt{\left(\sum_{l=1}^{N} w_l X_{il}^2 - \left(\sum_{l=1}^{N} w_l X_{il}\right)^2 \Big/ W_{ij}\right)\left(\sum_{l=1}^{N} w_l X_{jl}^2 - \left(\sum_{l=1}^{N} w_l X_{jl}\right)^2 \Big/ W_{ij}\right)}}$$

Noncomputable coefficients are set to system missing. The significance level for r_{ij} is based on

$$t = r_{ij}\sqrt{\frac{W_{ij} - 2}{1 - r_{ij}^2}}$$

which, under the null hypothesis, is distributed as a t with $W_{ij} - 2$ degrees of freedom. By default, one-tailed significance levels are printed.

Means and Standard Deviations

$$\overline{X}_j = \sum_{i=1}^{N} w_i X_{ji} / W_j$$

$$S_j = \sqrt{\left(\sum_{i=1}^{N} w_i X_{ji}^2 - \overline{X}_j^2 W_j \right) \Big/ \left(W_j - 1 \right)}$$

If pairwise deletion is selected, means and standard deviations are based on *all* nonmissing cases. For listwise deletion, only cases with no missing values on any specified variables are included.

Partial Correlations

Partial correlations are calculated recursively from the lower-order coefficients using

$$r_{ij.k} = \frac{r_{ij} - r_{ki} r_{kj}}{\sqrt{\left(1 - r_{ki}^2\right)\left(1 - r_{kj}^2\right)}} \quad \text{(first order)}$$

$$r_{ij.kl} = \frac{r_{ij.k} - r_{il.k} r_{jl.k}}{\sqrt{\left(1 - r_{il.k}^2\right)\left(1 - r_{jl.k}^2\right)}} \quad \text{(second order)}$$

and similarly for higher orders (Morrison 1976, p. 94).

If the denominator is less than 10^{-20}, or if any of the lower-order coefficients necessary for calculations are system missing, the coefficient is set to system missing. If a coefficient in absolute value is greater than 1, it is set to system missing. (This may occur with pairwise deletion.)

Significance Level

The significance level is based on

$$t = r\sqrt{\frac{df}{1-r^2}}$$

The degrees of freedom are

$$df = M - \theta - 2$$

where θ is the order of the coefficient and *M* is the minimum sum of weights from which the zero-order coefficients involved in the computations were calculated. Thus, for $r_{ij.kl}$

$$M = \min\left(W_{ij}, W_{ki}, W_{kj}, W_{il}, W_{lk}, W_{jl}\right)$$

where W_{ij} is the sum of weights of the cases used to calculated r_{ij}. If listwise deletion of missing values (default) was used, all W_{ij} are equal. By default, the significance level is one-tailed.

Reference

Morrison, D. F. 1976. *Multivariate statistical methods*. New York: McGraw-Hill.