Partial correlation

In the case of multivariate data say, $x_1, x_2, \cdots, x_p$ we may be interested to study of the degree to which two variables say x_1 and x_2 may be said to be related when the influence of the other $(p-2)$ variables $x_3, x_4, \cdots, x_p$ is eliminated from both of them.

Now if $X_{1\cdot 34\cdots p}$ and $X_{2\cdot 34\cdots p}$ are the estimated value of X_1 and X_2 respectively and the errors are

$$e_{1\cdot 34\cdots p} = X_1 - X_{1\cdot 34\cdots p}$$

$$\text{and } e_{2\cdot 34\cdots p} = X_2 - X_{2\cdot 34\cdots p}$$

then the product moment correlation between $e_{1\cdot 34\cdots p}$ and $e_{2\cdot 34\cdots p}$ is known as the partial correlation between X_1 and X_2 and is denoted by,

$$r_{12\cdot 34\cdots p} = \frac{\operatorname{cov}(e_{1\cdot 34\cdots p}\, e_{2\cdot 34\cdots p})}{\sqrt{\operatorname{Var}(e_{1\cdot 34\cdots p})\operatorname{Var}(e_{2\cdot 34\cdots p})}}$$

Here we have

$$\left.\begin{aligned} \cancel{E}\frac{1}{n}\sum(e_{1\cdot 34\cdots p}) &= \cancel{E}\frac{1}{n}\sum(e_{2\cdot 34\cdots p}) = 0 \\ \frac{1}{n}\sum(X_i\, e_{1\cdot 34\cdots p}) &= \frac{1}{n}\sum(X_i\, e_{2\cdot 34\cdots p}) = 0 \end{aligned}\right\} \text{ (*)} \quad \text{for } i = 3, 4, \ldots$$

[from least square equations

(*) where sum is taken over all 'i' possible values of the variables]

$$\operatorname{Var}(e_{1\cdot 34\cdots p}) = \sigma_1^2 \frac{R^{(2)}}{R_{11}^{(2)}}$$

$$\text{and } \operatorname{Var}(e_{2\cdot 34\cdots p}) = \sigma_2^2 \frac{R^{(1)}}{R_{22}^{(1)}}$$

Where $R^{(1)}$ is the determinant obtain from

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{pmatrix}_{p \times p}$$

by eliminating the 1st row and 1st column. While $R_{ij}^{(1)}$ is the cofactor of r_{ij} in $R^{(1)}$ and $R_{ij}^{(2)}$ are similarly define.

Now, $Cov(e_{1\cdot 34\cdots p}, e_{2\cdot 34\cdots p})$

$$= \frac{1}{n}\sum(e_{1\cdot 34\cdots p} \cdot e_{2\cdot 34\cdots p})$$

$$= \frac{1}{n}\sum\left\{(x_1-\bar{x}_1) + \frac{\sigma_1}{\sigma_3}\frac{R_{13}^{(2)}}{R_{11}^{(2)}}(x_3-\bar{x}_3) + \cdots + \frac{\sigma_1}{\sigma_p}\frac{R_{1p}^{(2)}}{R_{11}^{(2)}}(\cdots \right. \left. \times e_{2\cdot 34\cdots p}\right\}$$

$$= \frac{1}{n}\sum\left[(x_1-\bar{x}_1)\, e_{2\cdot 34\cdots p}\right] \quad [\text{from } (*)]$$

$$= \frac{1}{n}\sum\left[(x_1-\bar{x}_1)\left\{(x_2-\bar{x}_2) + \frac{\sigma_2}{\sigma_3}\frac{R_{23}^{(1)}}{R_{22}}(x_3-\bar{x}_3) + \cdots + \frac{\sigma_2}{\sigma_p}\frac{R_{2p}^{(1)}}{R_{22}^{(1)}} \cdots\right.\right.$$

$$= r_{12}\sigma_1\sigma_2 + \frac{R_{23}^{(1)}}{R_{22}^{(1)}} r_{13}\sigma_1\sigma_2 + \cdots + \frac{R_{2p}^{(1)}}{R_{22}^{(1)}} r_{1p}\sigma_1\sigma_2$$

$$= \frac{\sigma_1\sigma_2}{R_{22}^{(1)}}\left[r_{12}R_{22}^{(1)} + r_{13}R_{23}^{(1)} + \cdots + r_{1p}R_{2p}^{(1)}\right]$$

Now, $r_{12}R_{22}^{(1)} + r_{13}R_{23}^{(1)} + \cdots + r_{1p}R_{2p}^{(1)}$ is the determinant obtained from $R^{(1)}$ by replacing the first row with $r_{12}, r_{13} \cdots r_{1p}$ i.e. (-1) th the cofactor of r_{12} i.e $-R_{12}$

Now since. $R^{(1)} = R_{11}$, $R^{(2)} = R_{22}$ and $R_{11}^{(2)} = R_{22}^{(1)}$

Hence $$r_{12\cdot 34\cdots p} = \frac{-\sigma_1\sigma_2 \dfrac{R_{12}}{R_{22}^{(1)}}}{\sigma_1\sigma_2\sqrt{\dfrac{R^{(2)}}{R_{22}^{(1)}} \times \dfrac{R^{(1)}}{R_{22}^{(1)}}}} = -\frac{R_{12}}{\sqrt{R_{22}R_{11}}}$$