THE TECHNIQUE OF PATH COEFFICIENTS

MAX D. ENGELHART

The Chicago City Junior Colleges

A derivation of equations fundamental to the technique of path coefficients is given. Suggestions are made with respect to the calculations required in the use of the technique. The relations of the technique to those of partial correlation, semi-partial correlation, part correlation, multiple correlation, and factor analysis are discussed. Some consideration is given to the merits and limitations of the technique of path coefficients.

The technique of path coefficients is the contribution of Sewall Wright, who reported his derivations in 1921. (9) The following derivation represents an adaptation based upon proofs of a number of equalities given in an article by Dunlap and Cureton. (3) In this treatment a variable considered to be an effect of several others is labeled the "dependent" variable. The variables regarded as causes are labeled the "independent" variables, even though they may not be statistically independent, i.e., uncorrelated.

Let us assume that all of the variance of $x_0$, a dependent variable, is due to the correlated independent variables $x_1$ and $x_2$. Let the element $a_{01.2}$ represent that part of variable $x_0$ determined by its regression on $x_1$; and let $a_{02.1}$ represent that part of $x_0$ determined by its regression on $x_2$. Then:

\[ \begin{align*}
\mathbf{t} \quad x_0 &= a_{01.2} + a_{02.1} \\
\frac{\Sigma x_0^2}{N} &= \frac{\Sigma (a_{01.2} + a_{02.1})^2}{N} \\
\frac{\Sigma x_0^2}{N} &= \frac{\Sigma a_{01.2}^2}{N} + \frac{\Sigma a_{02.1}^2}{N} + \frac{2 \Sigma a_{01.2} a_{02.1}}{N}
\end{align*} \]

(1)

\[ \mathbf{t} \text{All variables are taken as measured from their respective means as origins.} \]
By definition*

\[ a_{01:2} = r_{0(1:2)} \frac{\sigma_0}{\sigma_{1:2}} x_1 \]

\[ a_{02:1} = r_{0(2:1)} \frac{\sigma_0}{\sigma_{2:1}} x_2 \]

\[ \frac{\Sigma a_{01:2}^2}{N} = r_{0(1:2)}^2 \frac{\sigma_0^2}{\sigma_{1:2}^2} \frac{\Sigma x_1^2}{N} = r_{0(1:2)}^2 \frac{\sigma_0^2}{\sigma_{1:2}^2} \]

\[ \frac{\Sigma a_{02:1}^2}{N} = r_{0(2:1)}^2 \frac{\sigma_0^2}{\sigma_{2:1}^2} \frac{\Sigma x_2^2}{N} = r_{0(2:1)}^2 \frac{\sigma_0^2}{\sigma_{2:1}^2} \]  \[ \tag{2} \]

\[ \frac{2 \Sigma a_{01:2} a_{02:1}}{N} = 2 r_{0(1:2)} \frac{\sigma_0}{\sigma_{1:2}} r_{0(2:1)} \frac{\sigma_0}{\sigma_{2:1}} \frac{\Sigma x_1 x_2}{N} \]

Since

\[ \frac{\Sigma x_1 x_2}{N} = \frac{\Sigma x_1 x_2}{N} \cdot \sigma_1 \sigma_2 = r_{12} \sigma_1 \sigma_2 \]

we have

\[ \frac{2 \Sigma a_{01:2} a_{02:1}}{N} = 2 r_{0(1:2)} \frac{\sigma_0}{\sigma_{1:2}} r_{0(2:1)} \frac{\sigma_0}{\sigma_{2:1}} \]

\[ = 2 \sigma_{a_{01:2}} \sigma_{a_{02:1}} r_{12} \]  \[ \tag{4} \]

Substituting from (2), (3), and (4) in (1) and writing \( \sigma_0^2 \), the variance of \( x_0 \), for \( \frac{\Sigma x_0^2}{N} \):

\[ \sigma_0^2 = \sigma_{a_{01:2}}^2 + \sigma_{a_{02:1}}^2 + 2 r_{12} \sigma_{a_{01:2}} \sigma_{a_{02:1}} . \]

*These equations are analogous to the ordinary regression equation for deviation measures, i.e., \( x_0 = r_{01} \frac{\sigma_0}{\sigma_1} x_1 \). Here, \( a_{01:2} \) represents a part of \( x_0 \); (determined by \( x_1 \) and independent of \( x_2 \)); \( r_{0(1:2)} \) represents the correlation between \( x_0 \) and that part of \( x_1 \) which is independent of \( x_2 \); and \( \sigma_{1:2} \) represents the corresponding partial standard deviation, i.e., the standard deviation of that part of \( x_1 \) which contributes to \( x_0 \) independently of \( x_2 \). \( r_{0(1:2)} \) and \( r_{0(2:1)} \) are coefficients of semi-partial correlation.