Multiple R^2 and Partial Correlation/Regression Coefficients[©]

b_i is an unstandardized partial slope.

Were we to predict Y from X_2 and predict X_1 from X_2 and then use the residuals from X_1, that is, $(X_1 - \hat{X}_{1\bullet 2})$, to predict the residuals in Y, that is, $(Y - \hat{Y}_2)$, the slope of the resulting regression would be b_1. That is, b_1 is the number of units that Y changes per unit change in X_1 after we have removed the effect of X_2 from both X_1 and Y. Put another way, b_i is the average change in Y per unit change in X_i with all other predictor variables held constant.

β is a standardized partial slope.

Look at the formulas for a trivariate multiple regression.

$$\hat{Z}_y = \beta_1 Z_1 + \beta_2 Z_2 \qquad \beta_1 = \frac{r_{y1} - r_{y2}r_{12}}{1 - r_{12}^2} \qquad \beta_2 = \frac{r_{y2} - r_{y1}r_{12}}{1 - r_{12}^2}$$

β_1 represents the unique contribution of X_1 towards predicting Y in the context of X_2. We must remove the effect of X_2 upon both X_1 and Y to obtain this unique contribution. If you look at the formula for β_1 you will see how this is done: The larger r_{y1}, the larger the β_1. Also, the larger the r_{y2} and the r_{12}, the smaller the β_1 (due to greater redundancy between X_1 and X_2 with respect to their overlap with Y).

Were we to predict Z_Y from Z_2, and Z_1 from Z_2, and then use the residuals from Z_1, that is, $(Z_1 - \hat{Z}_{1\bullet 2})$, to predict the residuals in Z_Y, that is, $(Z_Y - \hat{Z}_{Y\bullet 2})$, the slope of the resulting regression would be β_1. That is, β_1 is the number of standard deviations that Y changes per standard deviation change in X_1 after we have removed the effect of X_2 from both X_1 and Y. It should be clear that the value of β_i can be greatly affected by the correlations of other predictors with Y and with X_i. The notation $\hat{X}_{1\bullet 23}$ stands for X_1 predicted from X_2 and X_3.

R^2 can be interpreted as a simple r^2, a proportion of variance explained.

$$R^2_{Y\bullet 12 \ldots i \ldots p} = r^2_{y\hat{y}} = \frac{\sigma^2_{\hat{y}}}{\sigma^2_y}$$

The variance in predicted Y, that is, $\sigma^2_{\hat{y}}$, represents differences in Y due to the linear "effect" upon Y of the optimally weighted combination of the X's.

Thus, R^2 represents the proportion of the total variance in Y that is explained by the linear relationship between Y and the weighted combination of X's.

R^2 can be obtained from beta weights and zero-order correlation coefficients.

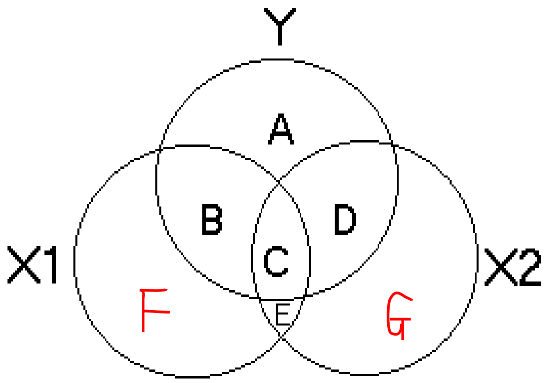
$$R^2_{Y\bullet 12\ldots i\ldots p} = \sum \beta_i r_{yi} = \sum \beta_i^2 + 2\sum \beta_i \beta_j r_{ij} \qquad (i \neq j)$$

The sum of the squared beta weights represents the sum of the unique contributions of the predictors while the rightmost term represents the redundancy among the predictors.

$$R^2_{y\bullet 12} = \frac{r^2_{Y1} + r^2_{Y2} - 2r_{Y1}r_{Y2}r_{12}}{1 - r^2_{12}} = \beta_1 r_{Y1} + \beta_2 r_{Y2}$$

Note that in determining R^2 we have added together the two bivariate coefficients of determination and then corrected (reduced) that sum for the redundancy of X_1 and X_2 in predicting Y.

Squared correlation coefficients represent proportions of variance explained.



$$a + b + c + d = 1$$

$$r^2_{Y1} = b + c$$

$$sr^2_1 = \frac{b}{(a+b+c+d)=1} = b$$

$$r^2_{Y2} = d + c$$

$$r^2_{12} = c + e$$

$$R^2_{Y\cdot 12} = b + c + d$$

c = redundancy (aka commonality)

A **squared semipartial correlation** represents the proportion of all the variance in Y that is associated with one predictor but not with any of the other predictors. That is, in terms of the Venn diagram,

$$sr_1^2 = \frac{b}{(a+b+c+d)=1} = b.$$

The squared semipartial can also be viewed as the decrease in R^2 that results from removing a predictor from the model, that is,

$$sr_i^2 = R^2_{Y\bullet 12\ldots i\ldots p} - R^2_{Y\bullet 12\ldots(i)\ldots p}$$

In terms of residuals, the **semipartial correlation** for X_i is the r between all of Y and X_i from which the effects of all other predictors have been removed. That is,

$$sr_1 = \text{corr between Y and } (X_1 - \hat{X}_{1\bullet 2})$$

A **squared partial correlation** represents a fully partialled proportion of the variance in Y: Of the variance in Y that is not associated with any other predictors, what proportion is associated with the variance in X_i. That is, in terms of the Venn diagram,

$$pr_1^2 = \frac{b}{a+b}$$

The squared partial can be obtained from the squared semipartial:

$$pr_i^2 = \frac{sr_i^2}{1 - R^2_{Y\bullet 12\ldots(i)\ldots p}} \qquad\qquad pr_i^2 > sr_i^2$$

The (i) in the subscript indicates that X_i is not included in the R^2.

In terms of residuals, the **partial correlation** for X_i is the r between Y from which all other predictors have been partialled and X_i from which all other predictors have been removed. That is,

$$pr_1 = \text{corr between } (Y - \hat{Y}_2) \text{ and } (X_1 - \hat{X}_{1\bullet 2})$$

If the predictors are well correlated with one another, their partial and semipartial coefficients may be considerably less impressive than their zero-order coefficients. In this case it might be helpful to conduct what some call a commonality analysis. In such an analysis one can determine how much of the variance in Y is related to the predictors but not included in the predictors partial or semipartial coefficients. For the Venn diagram above, that is area c. For more details, please see my document Commonality Analysis.

Return to Wuensch's Stats Lessons Page