# Review of: "Decoding the Correlation Coefficient: A Window into Association, Fit, and Prediction in Linear Bivariate Relationships"

Szabolcs Blazsek[1]

1 Universidad Francisco Marroquín

Potential competing interests: No potential competing interests to declare.

Review of “Decoding the Correlation Coefficient: A Window into Association, Fit, and Prediction in Linear Bivariate Relationships”

The paper in general is well-written, clear, and easy to read. I believe that the discussion of the methods about the difference between the correlation coefficient of the dependent and explanatory variables in a simple linear regression and the parameter estimate for the explanatory variable is well-known in the literature. So, I believe this paper can be considered as a review article, rather than an original research article.

Comment 1. "In a linear bivariate relationship, the coefficient and the regression slope are not completely independent, and thus we cannot assert with certainty that a higher correlation coefficient (r) never implies a stronger correspondence in change"

I think that word is missing from this sentence: "In a linear bivariate relationship, the CORRELATION coefficient and the regression slope are not completely independent, and thus we cannot assert with certainty that a higher correlation coefficient (r) never implies a stronger correspondence in change." I have the same comment for the sentence that follows this sentence in the paper.

Comment 2. The author may include the following equations in the paper to improve the explanation of the differences between the correlation coefficient and \beta_{1} for a simple linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\beta_1 = \frac{\mathrm{Cov}(y_i, x_i)}{\sigma_x^2}$$

$$\rho_{x,y} = \frac{\mathrm{Cov}(y_i, x_i)}{\sigma_y \sigma_x}$$

Relationship between the correlation coefficient and beta_{1}:

$$\beta_1 \frac{\sigma_x}{\sigma_y} = \frac{\mathrm{Cov}(y_i, x_i)}{\sigma_y \sigma_x} = \rho_{x,y}$$

Coefficient of determination:

$$R^2 = 1 - \frac{\sigma_\epsilon^2}{\sigma_y^2}$$

But I think that the coefficient of determination is not needed to be mentioned in this paper.

I think that when the differences between \beta_{1} and \rho_{x,y} is discussed it is useful to mention that, under the exogeneity assumption for x_{i}, \beta_{1} can be interpreted as a measure of causality which provides forecasts of y_{i} for the user of the linear regression model. The correlation coefficient does not measure causality or forecasting relation between y_{i} and x_{i}, but the association between the two variables.

Comment 3.

“Joseph Lee Rodgers and W. Alan Nicewander (1988) in their article ”Thirteen Ways to Look at the Correlation Coefficient"" could be replaced by Rodgers and Nicewander (1998).