How Do You Interpret Your Regression Coefficients?

Vijayamohanan Pillai N.

2016

How Do You Interpret Your Regression Coefficients?

Vijayamohanan Pillai N.

Centre for Development Studies, Kerala, India.

e-mail; vijayamohan@cds.ac.in

Abstract

This note is in response to David C. Hoaglin's provocative statement in *The Stata Journal* (2016) that "Regressions are commonly misinterpreted". "Citing the preliminary edition of Tukey's classic *Exploratory Data Analysis* (1970, chap. 23), Hoaglin argues that *the* correct interpretation of a regression coefficient is that it "tells us how *Y* responds to change in *X*2 after adjusting for simultaneous linear change in the other predictors in the data at hand". He contrasts this with what he views as the common misinterpretation of the coefficient as "the average change in *Y* for a 1-unit increase in *X*2 when the other *X*s are held constant". He asserts that this interpretation is incorrect because "[i]t does not accurately reflect how multiple regression works". We find that Hoaglin's characterization of common practice is often inaccurate and that his narrow view of proper interpretation is too limiting to fully exploit the potential of regression models. His article rehashes debates that were settled long ago, confuses the estimator of an effect with what is estimated, ignores modern approaches, and rejects a basic goal of applied research." (Long and Drukker, 2016:25). This note broadly agrees with the comments that followed his article in the same issue of *The Stata Journal* (2016) and seeks to present an argument in favour of the commonly held interpretation that Hoaglin unfortunately marks as misinterpretation.

How Do You Interpret Your Regression Coefficients?

Vijayamohanan Pillai N.

This note is in response to David C. Hoaglin's provocative statement in *The Stata Journal* (2016) that "Regressions are commonly misinterpreted". This note broadly agrees with the comments that followed his article in the same issue of *The Stata Journal* (2016) and seeks to present an argument in favour of the commonly held interpretation that Hoaglin unfortunately marks as misinterpretation. His argument was succinctly presented by J. Scott Long and David M. Drukker along with their comments as follows:

"Citing the preliminary edition of Tukey's classic *Exploratory Data Analysis* (1970, chap. 23), Hoaglin argues that *the* correct interpretation of a regression coefficient is that it "tells us how *Y* responds to change in *X*2 after adjusting for simultaneous linear change in the other predictors in the data at hand". He contrasts this with what he views as the common misinterpretation of the coefficient as "the average change in *Y* for a 1-unit increase in *X*2 when the other *X*s are held constant". He asserts that this interpretation is incorrect because "[i]t does not accurately reflect how multiple regression works". We find that Hoaglin's characterization of common practice is often inaccurate and that his narrow view of proper interpretation is too limiting to fully exploit the potential of regression models. His article rehashes debates that were settled long ago, confuses the estimator of an effect with what is estimated, ignores modern approaches, and rejects a basic goal of applied research." (Long and Drukker, 2016:25).

The present note is sympathetic to the above arguments, and attempts to substantiate the classical (text book) interpretation using the concept of the partial correlation coefficient.

We start with the text book interpretation by considering the following multiple regression with two explanatory variables, X_1 and X_2:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i \text{ ; } i = 1, 2, \ldots, N. \qquad \ldots.\ (1)$$

According to the text book interpretation, X_1 is said to be the covariate with respect to X_2 and vice versa. Covariates act as controlling factors for the variable under consideration. In the presence of the control variables, the regression coefficients βs are partial regression coefficients. Thus, β_1 represents the marginal effect of X_1 on Y, keeping all other variables, here X_2, constant. The latter part, that is, keeping X_2 constant, means the marginal effect of X_1 on Y is obtained after removing the linear effect of X_2 from *both* X_1 and Y. A similar explanation goes for β_2 also. Thus multiple regression facilitates to obtain the *pure* or *net* marginal effects by including all the relevant covariates and thus controlling for their heterogeneity.

This we'll discuss in a little detail below. We begin with the concept of partial correlation coefficient. Suppose we have three variables, X_1, X_2 and X_3. The simple correlation coefficient r_{12} gives the degree of correlation between X_1 and X_2. It is possible that X_3 may have an influence on both X_1 and X_2. Hence a question comes up: Is an observed correlation between X_1 and X_2 merely due to the influence of X_3 on both? That is, is the correlation merely due to the common influence of X_3? Or, is there a *net* correlation between X_1 and X_2, over and above the correlation due to the common influence of X_3? It is this *net* correlation between X_1 and X_2 that the partial correlation coefficient captures after removing the influence of X_3 from each, and then estimating the correlation between the *unexplained* residuals that remain. To prove this, we define the following:

Coefficients of correlation between X_1 and X_2, X_1 and X_3, and X_2 and X_3 are given by r_{12}, r_{13}, and r_{23} respectively, defined as

$$r_{12} = \frac{\sum x_1 x_2}{\sqrt{\sum x_1^2 \sum x_2^2}} = \frac{\sum x_1 x_2}{s_1 s_2},\ r_{13} = \frac{\sum x_1 x_3}{\sqrt{\sum x_1^2 \sum x_3^2}} = \frac{\sum x_1 x_3}{s_1 s_3} \text{ and } r_{23} = \frac{\sum x_2 x_3}{\sqrt{\sum x_2^2 \sum x_3^2}} = \frac{\sum x_2 x_3}{s_2 s_3}. \ldots(2')$$

Note that the lower case letters, x_1, x_2, and x_3, denote the respective variables in mean deviation (or demeaned) form; thus $(x_1 = X_{1i} - \bar{X}_1)$, etc. Thus, for example, $\Sigma x_1 x_2$ gives the covariance of X_1 and X_2 and Σx_1^2, the variance of X_1, the square root of which is its standard deviation (SD), such that s_1, s_2, and s_3 denote the SDs of the three variables.

The common influence of X_3 on both X_1 and X_2 may be modeled in terms of regressions of X_1 on X_3, and X_2 on X_3, with b_{13} as the slope of the regression of X_1 on X_3, given (in deviation form) by $b_{13} = \frac{\sum x_1 x_3}{\sum x_3^2} = r_{13}\frac{s_1}{s_3}$, and b_{23} as that of the regression of X_2 on X_3 given by $b_{23} = \frac{\sum x_2 x_3}{\sum x_3^2} = r_{23}\frac{s_2}{s_3}$.

Given these regressions, we can find the respective unexplained residuals. The residual from the regression of X_1 on X_3 (in deviation form) is $e_{1.3} = x_1 - b_{13} x_3$, and that from the regression of X_2 on X_3 is $e_{2.3} = x_2 - b_{23} x_3$.

Now the partial correlation between X_1 and X_2, net of the effect of X_3, denoted by $r_{12.3}$, is defined as the correlation between these *unexplained* residuals and is given by $r_{12.3} = \frac{\sum e_{1.3} e_{2.3}}{\sqrt{\sum e_{1.3}^2}\sqrt{\sum e_{2.3}^2}}$. Note that since the least-squares residuals have zero means, we need not write them in mean deviation form. We can directly estimate the two sets of residuals and then find out the correlation coefficient between them. However, the usual practice is to express them in terms of simple correlation coefficients. Using the definitions given above of the residuals and the regression coefficients, we have for the residuals: $e_{1.3} = x_1 - r_{13}\frac{s_1}{s_3}x_3$, and $e_{2.3} = x_2 - r_{23}\frac{s_2}{s_3}x_3$, and hence, upon simplification, we get

$$r_{12.3} = \frac{\sum e_{1.3}e_{2.3}}{\sqrt{\sum e_{1.3}^2}\sqrt{\sum e_{2.3}^2}} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}.$$

"This is the statistical equivalent of the economic theorist's technique of impounding certain variables in a *ceteris paribus* clause." (Johnston, 1972: 58). Thus the partial correlation coefficient between X_1 and X_2 is said to be obtained by keeping X_3 constant. This idea is clear in the above formula for the partial correlation coefficient as a *net* correlation between X_1 and X_2 after removing the influence of X_3 from each.

When this idea is extended to multiple regression coefficients, we have the partial derivatives as the partial regression coefficients. Consider the regression equation in three variables, X_1, X_2 and X_3:

$$X_{1i} = \alpha + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \,;\; i = 1, 2, \ldots, N. \qquad \ldots. (3)$$

Since the estimated regression coefficients are partial ones, the equation can be written as:

$$X_{1i} = a + b_{12.3}X_{2i} + b_{13.2}X_{3i}\,, \qquad \ldots. (4)$$

where the lower case letters (a and b) are the OLS estimates of α and β respectively.

The estimate $b_{12.3}$ is given by:

$$b_{12.3} = \frac{\sum x_1x_2 \sum x_3^2 - \sum x_1x_3 \sum x_2x_3}{\sum x_2^2 \sum x_3^2 - (\sum x_2x_3)^2}.$$

Now using the definitions of simple and partial correlation coefficients in (2) and (2'), we can rewrite the above as:

$$b_{12.3} = \frac{r_{12} - r_{13} r_{23}}{1 - r_{23}^2} \frac{s_1}{s_2}.$$

Why $b_{12.3}$ is called a partial regression coefficient is now clear from the above definition: it is obtained after removing the common influence of X_3 from both X_1 and X_2.

Similarly, we have the estimate $b_{13.2}$ given by:

$$b_{13.2} = \frac{\sum x_1 x_3 \sum x_2^2 - \sum x_1 x_2 \sum x_2 x_3}{\sum x_2^2 \sum x_3^2 - (\sum x_2 x_3)^2} = \frac{r_{13} - r_{12} r_{32}}{1 - r_{23}^2} \frac{s_1}{s_3},$$

obtained after removing the common influence of X_2 from both X_1 and X_3.

Thus the fundamental idea in partial (correlation/regression) coefficient is estimating the *net* correlation between X_1 and X_2 after removing the influence of X_3 from each, by computing the correlation between the *unexplained* residuals that remain (after eliminating the influence of X_3 from both X_1 and X_2). The classical text books describe this procedure as controlling for or accounting for the effect of X_3, or keeping that variable constant; whereas Tukey characterizes this as "adjusting for simultaneous linear change in the other predictor", that is, X_3. Above all these seeming semantic differences, let us keep the underlying idea alive, while interpreting the regression coefficients.

References

Hoaglin, David C. (2016) "Regressions are commonly misinterpreted". *The Stata Journal* Vol. 16, Number 1, pp. 5–22.

Johnston, J. (1972) *Econometric Methods*. Second edition. McGraw-Hill.

Long, J. Scott and Drukker, David M. (2016) "Regressions are commonly misinterpreted: Comments on the article". *The Stata Journal* Vol. 16, Number 1, pp. 25–29.

Davidson, R. and J. G. MacKinnon. 2004. *Econometric theory and methods.* New York: Oxford University Press.

Frisch, R. and F.V.Waugh. 1933. .Partial time regression as compared with individual trends. *Econometrica* 1 (October): 387-401.

Green, W. H. 2003. *Econometric Analysis.* 5th ed., Upper Saddle River: Prentice Hall.
Johnston, J. and J. Dinardo.1997. *Econometric methods. 4th ed.* New York: McGraw Hill/Irwin.

Lovell, M. C. 1963. Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association* 58 (December): 993-l0l0.

Goldberger, A., 1991, A Course in Econometrics, Harvard University Press, Cambridge.

Davidson, R. and MacKinnon, R., 1993, Estimation and Inference in Econometrics, Oxford University Press, Oxford.

Fiebig, D., Bartels, R. and Kramer, W., 1996, The Frisch-Waugh Theorem and Generalized Least Squares, Econometric Reviews, 15(4), pp. 431-443.

Ruud, P., 2000, An Introduction to Classical Econometric Theory, Oxford University Press, Oxford.