How to interprete the results of regression models for program impact analysis

Kaz Uekawa

August 2020

The purpose of this memo is to describe how you can interpret coefficients from regression models and make conclusions about program impact. I will use a hypothetical study where we have data from 10,000 students who come from 40 schools. I will describe two experimental setups:

Scenario A: Schools are either treatment schools or comparison schools (or control schools).

Scenario B: Schools are either comparison school, treatment A schools, and treatment B schools.

For each of the scenario, we get results like the following. I made up the numbers.

Scenario A

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimates | Standard Errors | p-value | Stat test |
| Intrercept | 1.234 | 0.23 | 0.001 | \*\* |
| Treatment | 2.345 | 0.32 | 0.03 | \* |
| Pretest | 3.456 | 0.45 | 0.03 |  |
| Male | 4.567 | 0.56 | 0.03 |  |

Notes to SAS users. I usually code treatment status as a numeric variable (0 if comparison; 1 if treatment) and do not include it in the class line (of, for example, proc glimmix). You can use a string variable too instead and include it in the class line. If the values are “treatment” and “comparison,” you will get a coefficient for comparison and you will just get “.” (dot) for treatment as the latter is alphabetically later (?) and serves as the omitted category (AKA “reference category”). You can manipulate which group becomes the reference category by changing the text value. I would use the numeric prefix: “(1) treatment” and “(2)comparison” to make the comparison group the reference category.

Scenario B

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimates | Standard Errors | p-value | Stat test |
| Intercept | 1.234 | 0.23 | 0.001 | \*\* |
| Treatment A | 2.345 | 0.32 | 0.03 | \* |
| Treatment B | 4.324 | 0.23 | 0.02 |  |
| Pretest | 3.456 | 0.45 | 0.03 |  |
| Male | 4.567 | 0.56 | 0.03 |  |

How to interpret scenario A

Scenario A is straight forward because there are only two groups. The impact coefficient is the one from Treatment variable, 2.345. It is statistically significant (I made up this result). This means that the treatment students did better than comparison students and the difference was 2.345. This is what you want to report. If the impact coefficient is negative, it means comparison students did better than treatment students.

When people hear the program impact coefficient, 2.345, they may not understand what it means (though if it is a state standard test (SOL test in the state of Virginia) and readers know the test product, 2.345 may be meaningful to them.

To assist interpretation, you want to report a standardized effect size. You can redo the same analysis by using the z-score version of the outcome scores. Alternatively, you can divide the program coefficient (2.345 in this example) by the outcome variable’s SD (you can use the sample SD or use the comparison group’s SD; you have to decide with a rational).

I will make up a number here. The standardized effect size was .25. You can understand this in terms of standard deviation in the outcome scores. We typically consider 0.25 as small effect in education research.

Now that you reported that the program impact was 2.345 and the standardized program impact was 0.25, you can use the graph for readers who like to see visuals. You can use a hand-calculator to derive the adjusted averages (I usually use SAS or R to do this calculation right after I run a regression model). The adjusted average of the comparison group is 1.234 (the intercept value). The treatment group’s adjusted average is 1.234+2.345 (=3.579). We can create a bar graph. You can call these “adjusted” averages because the multivariate models used multiple predictors when deriving the program impact coefficient.

Sorry, I will completely “delete” what I just said above. The meaning of 1.23 is unclear as is because it came from the intercept value of the regression model. You might have noticed that the intercept value feels impossible to interpret. It is the average value for a subject (in the sample) whose predictor values are all 0s. For example, if you have a gender as a predictor and the values are coded as 1 if boy and 0 if girl, the intercept value represents the average for girls. One tricky thing is that in this example case, I have a pretest score variable in it. Imagine its score distribution starts from 300 and ends with 500. There is no subject whose pretest score is 0. Thus, the intercept value of 1.23 is artificial and non-sensical.

One approach, which I don’t do any more, is to make the pretest variable or all continuous variables in the predictor Z-scores. However, this is extra work and runs a risk of making errors.

This is what I suggest you do. Just use the simple average mean of the comparison group for the comparison group value. Let me make up a number, 1.35. Use that as the comparison group’s mean. Because the treatment effect (program impact coefficient) is 2.345, the adjusted mean for the treatment group (for the graphic purpose) is 1.35+ 2.345 (=3.70).

Notes: The unadjusted comparison group mean is used for the comparison group mean. The treatment group mean is the adjusted one from the regression model.

“Adjusted” means the program impact estimate is adjusted for multiple predictors used in the regression model.

To conclude, you want to report:

1. The program impact coefficient (AKA program effect)
2. The standardized version of the program impact coefficient (AKA standardized effect)
3. a bar graph showing the group means for the two groups (the comparison group’s mean comes from a simple descriptive statistics; the treatment groups’ mean = that + the program impact coefficient).

How to interpret scenario B

Let’s take a look at the same table again.

Scenario B

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimates | Standard Errors | p-value | Stat test |
| Intrecept | 1.234 | 0.23 | 0.001 | \*\* |
| Treatment A | 2.345 | 0.32 | 0.03 | \* |
| Treatment B | 4.324 | 0.23 | 0.02 |  |
| Pretest | 3.456 | 0.45 | 0.03 |  |
| Male | 4.567 | 0.56 | 0.03 |  |

This means that the program impact for Treatment A is 2.345. The treatment effect for Treatment B is 4.324. In other words, treatment A students did better than comparison students and the difference was 2.345. Treatment B students did better than comparison students and the difference was 4.324.

You may be also interested in evaluating the difference between Treatment A and B. You can do this by treating either treatment A or B as the omitted category (reference group).

Table above came from a regression model like this. The omitted category (reference group) was the comparison group.

Postest score = intercept +b1\*treatmentA + b2\*tretmentB +b3\*Pretest + b4\*Male

If you can do this to make Treatment B as the committed category. The omitted category is the treatment B.

Postest score = intercept +b1\*treatmentA + b2\*comparison +b3\*Pretest + b4\*Male

I made up numbers. Now you can tell that the difference between treatment A and B is 1.345. You can also say the difference is statistically significant (\*).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimates | Standard Errors | p-value | Stat test |
| Intercept | 3.234 | 0.23 | 0.001 | \*\* |
| Treatment A | 1.345 | 0.32 | 0.03 | \* |
| Comparison | 3.324 | 0.23 | 0.02 |  |
| Pretest | 3.456 | 0.45 | 0.03 |  |
| Male | 4.567 | 0.56 | 0.03 |  |

So by changing the omitted category, you can see how the three groups are different from one another.

Logistic Regression Model

The discussion so far applies to both linear regression models (e.g., OLS regression) and logistic regression model. How to interpret coefficients do not change if the model is a multilevel model or a fixed effect model.

If logistic regression model, however, coefficients are not very intuitive. For example, what does 1.23 mean? Coefficients from the logistic regression models are expressed in logit. What does a logit value of 1.23 mean?

I usually convert the logit into odds ratio and probabilities. I have an Excel sheet that does the conversion of logit values into odds ratio and probabilities (of two groups).

<https://drive.google.com/file/d/0B7AoA5fyqX_sN0RUc0E5aFowb00/view>