Statistical Modeling and Causal Inference with R

Week 10: Moderation and heterogeneous effects

Manuel Bosancianu

Max Schaub

November 16, 2020

Hertie School of Governance

Lecture Q&A

- ✓ Open Q&A
- ✓ Remarks on assignments
- ✓ Quiz interpreting regression coefficients

Remarks on assignments

1. Don't leave working on the assignment until shortly before the deadline – this makes it impossible to get help; when running into serious problems (e.g. getting stuck for a couple of hours), seek help in the group or from the TAs
2. Find a good 'balance-table' routine from the tutorials by Adelaida and Sebastian, and use it consistently – balance tables are very useful (and will be asked for in the exam)
3. Do read the question wording very carefully – often, points are lost because you do something that's not been asked for.
4. Technical note: Turn off scientific notification in R since this seems to lead to confusion; turn it off (options(scipen=999))

Quiz interpreting regression coefficients

What you need:

- ✓ A calculator (or being quick at mental arithmetic)
- ✓ Knowledge of the lecture slides

Instructions:

- ✓ Calculate the response
- ✓ Post your answer in the poll at *https://www.menti.com/tjphg6xova*

Quiz interpreting regression coefficients

You are estimating the effect of a mentoring program on income.

For the evaluation 185 individuals were randomly assigned to a treatment group that received mentoring, while 185 were assigned to a control group that did not.

You now have data on their incomes 15 months after the intervention. In order to estimate the effect, you are using OLS regression models.

You start with the simple model in the form:

Model 1: $Y_i = \beta_0 + \beta_1 D_i + \mu_i$

Where Y_i stands for program participants' income and D_i for the mentoring program.

Quiz interpreting regression coefficients

The results are shown here:

Model: $Y_i = \beta_0 + \beta_1 D_i + \mu_i$

	Income
Mentoring	2213**
	(717)
Intercept	4574**
	(507)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Q1: What is the treatment effect, i.e. the partial effect $\frac{\partial Y_i}{\partial D_i}$ of the mentoring program?

Answer: $\beta_1 = 2213$

Quiz interpreting regression coefficients

Model: $Y_i = \beta_0 + \beta_1 D_i + \mu_i$

	Income
Mentoring	2213**
	(717)
Intercept	4574**
	(507)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Q2: What is the income of those not in the mentoring program $E[Y_i|D_i = 0]$?

Answer: $\beta_0 = 4574$

Quiz interpreting regression coefficients

Model: $Y_i = \beta_0 + \beta_1 D_i + \mu_i$

	Income
Mentoring	2213**
	(717)
Intercept	4574**
	(507)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Q3: What is the average income for the whole sample $E[Y_i]$?

Answer: Weighted average

$\frac{185}{370} \times \beta_0 + \frac{185}{370} \times (\beta_0 + \beta_1) = 0.5 * 4574 + 0.5 * (4574 + 2213) = 5680.5$

Quiz interpreting regression coefficients

Now to improve the efficiency of your estimates, you add an indicator recording whether the participants have a Masters degree.

This is the indicator U that takes the value 1 for those with a Masters degree (87/370), and 0 for those without (283/370).

This means your second model takes the form:

Model 2: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 U_i + \epsilon_i$

Quiz interpreting regression coefficients

Model: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 U_i + \epsilon_i$

	Income
Mentoring	1962**
	(718)
Masters degree	2216**
	(847)
Intercept	4178**
	(525)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Q4: What is the income of those in the mentoring program but without a Masters degree $E[Y_i|D_i = 1, U_i = 0]$?

Answer: $\beta 0 + \beta 1 = 4178 + 1962 = 6140$

Quiz interpreting regression coefficients

You suspect that the mentoring program has a different effect for those with a Masters degree vs. those without. You therefore estimate an interaction model in the form:

Model 3: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 U_i + \beta_3 D_i \times U_i + \mu_i$

Quiz interpreting regression coefficients

Model: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 U_i + \beta_3 D_i \times U_i + \mu_i$

	Income
Mentoring	1048
	(811)
Masters degree	-156
	(1306)
Mentoring × Masters degree	4053*
	(1708)
Intercept	4601**
	(552)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Q5: What is the effect of the treatment for those without a Masters degree $\frac{\partial Y_i}{\partial D_i, U_i = 0}$?

Answer: $\beta 1 = 1048$

Quiz interpreting regression coefficients

Model: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 U_i + \beta_3 D_i \times U_i + \mu_i$

	Income
Mentoring	1048
	(811)
Masters degree	-156
	(1306)
Mentoring × Masters degree	4053*
	(1708)
Intercept	4601**
	(552)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Q6: What is the effect of the treatment for those **with** a Masters degree $\frac{\partial Y_i}{\partial D_i, U_i = 1}$? And, Q7, what is the difference between the treatment effects?

Answer Q6: $\beta 1 + \beta 3 = 1048 + 4053 = 5101$, Answer Q7: $\beta 3 = 4053, p < 0.01$ as per regression table.

Quiz interpreting regression coefficients

Model: $Y_i = \beta_0 + \beta_1 D_i + \beta_2 U_i + \beta_3 D_i \times U_i + \mu_i$

	Income
Mentoring	1048
	(811)
Masters degree	-156
	(1306)
Mentoring × Masters degree	4053*
	(1708)
Intercept	4601**
	(552)
N	370

Standard errors in parentheses. † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Remember that 87/370 participants have a Masters degree, and 283/370 don't.

Q8: What is the **overall** effect of the mentoring program, i.e. the partial effect $\frac{\partial Y_i}{\partial D_i}$?

Answer: Weighted average $\frac{283}{370}\beta_1 + \frac{87}{370} \times (\beta_1 + \beta_3) = 1048 * (283/370) + 5101 * (87/370) = 2001$

Thank you for watching, and see you next Monday!