Handout for Lecture 19

Categorical Variables & Interaction Terms

ECON 340: Economic Research Methods — Instructor: Div Bhagia

1. Consider the following regression model:

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 D \cdot X + u$$

Here, X represents a continuous variable, and D is a dummy variable that takes values 1 or 0. Assume that both X and D are exogenous. Write down the expressions for the following expectations.

$$E(Y|D = 1, X) = (\beta_0 + \beta_1) + (\beta_2 + \beta_3)X$$

$$E(Y|D = 0, X) = \beta_0 + \beta_2 X$$

What is the impact of changing D from 1 to 0 on Y? Does this impact vary by X?

$$E(Y|D = 1, X) - E(Y|D = 0, X) = \beta_1 + \beta_3 X$$

Yes, varies by X.

2. Consider the following regression model:

$$wages = \beta_0 + \beta_1 Female + \beta_2 Hispanic + \beta_3 Female \times Hispanic + u$$

Here, $Female$ is a dummy variable assigned the value of 1 if an individual's gender is female and 0 if not. Similarly, $Hispanic$ is a dummy variable that is set to 1 if an individual's ethnicity is Hispanic and 0 otherwise. The regression output for this model is given below. Answer the following questions.

	Wages
Intercept	70,179.09***
	(473.52)
Female	−16,046.81***
	(683.42)
Hispanic	−19,367.71***
	(1,211.46)
Female X Hispanic	8,163.75***
	(1,788.04)
Observations	17,578
R^2	0.05
Note:	*p<0.1; **p<0.05; ***p<0.01

(a) What is the average wage income for non-Hispanic males in this sample?

$$E(wages|Female = 0, Hispanic = 0) = \$70{,}179.09$$

(b) What is the average wage income for Hispanic males in this sample?

$$E(wages|Female = 0, Hispanic = 1) = 70{,}179.09\text{-}19{,}367.71 = \$50{,}811.38$$

(c) What is the average wage income for non-Hispanic females in this sample?

$$E(wages|Female = 1, Hispanic = 0) = 70{,}179.09\text{-}16{,}046.81 = \$54{,}132.28$$

(d) What is the average wage income for Hispanic females in this sample?

$$E(wages|Female = 1, Hispanic = 1) = 70{,}179.09\text{-}19{,}367.71\text{-}16{,}046.81\text{+}8{,}163.75$$
$$= \$42{,}928.32$$

(e) How do we interpret the coefficient on the interaction between $Hispanic$ and $Female$?

It is the difference between the gender-wage gap for Hispanic vs non-Hispanic individuals.

$$E(wages|Female = 0, Hispanic = 1) - E(wages|Female = 1, Hispanic = 1)$$
$$= 50811.38 - 42928.32 = 7883.06$$
$$E(wages|Female = 0, Hispanic = 0) - E(wages|Female = 1, Hispanic = 0)$$
$$= 70179.09 - 54132.28 = 16046.8$$

Difference between the two: 16046.8-7883.06 = \$8,163.73

Alternatively, can interpret β_3 as the difference in the impact of being hispanic by gender.