Discussion 5

Week of 02.16

UVA Batten

Agenda

- Quiz review
- More regression mechanics!

Announcements

- Quiz 4 due Tuesday, Feb. 17 at 9pm.
- Midterm exam on Wednesday February 25th
- Discussion sections next week will serve as extra office hours for the exam
- No quiz next week

Preparing for the Exam

In order of importance, we recommend using the following materials:

1. Lectures
2. Homeworks
3. Quizzes
4. Worksheets

As mentioned in the announcements, office hours, tutoring sessions, and class next week will serve as additional resources and review sessions.

TA Tip: Prof. Tello-Trillo values conceptual understanding. It is helpful if you feel comfortable being able to teach RMDA to someone with no stats experience.

Quiz Review

Question 2

A city evaluates the effects of a voluntary after-school tutoring program on math test scores with the following regression: $Score_i = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1 Tutoring_i$.

If $Var(Tutoring_i) = 0$, this would imply that $\boldsymbol{\alpha}_1 = 0$.

( ) True

( ) False

Question 2

A city evaluates the effects of a voluntary after-school tutoring program on math test scores with the following regression: $Score_i = \alpha_0 + \alpha_1 Tutoring_i$.

If $Var(Tutoring_i) = 0$, this would imply that $\alpha_1 = 0$.

( ) True

(✅) False

Question 3

A city evaluates the effects of a voluntary after-school tutoring program on math test scores with the following regression: $Score_i = \alpha_0 + \alpha_1 Tutoring_i$.

If $Corr(Tutoring_i\ , Score_i) = 0$, this would imply that $\alpha_1 = 0$.

( ) True

( ) False

Question 3

A city evaluates the effects of a voluntary after-school tutoring program on math test scores with the following regression: $Score_i = \alpha_0 + \alpha_1 Tutoring_i$.

If $Corr(Tutoring_i\ , Score_i) = 0$, this would imply that $\alpha_1 = 0$.

(✅) True

( ) False

Regression Mechanics

🅿️ Changing Units

Imagine you run the following regression on the relationship between secondary education (in years) and monthly earnings (in dollars).

$$Earnings = \alpha_0 + \alpha_1 Education$$

You find the value of α_0 to be 2000 and the value of α_1 to be 900. How would you interpret these values?

Now imagine you change the measurement of earnings to hundreds of dollars. How would this change your estimates of α_0 and α_1?

What about if you changed the measurement of education to months?

🅿️ Changing Units

Imagine you run the following regression on the relationship between secondary education (in years) and monthly earnings (in dollars).

$\boldsymbol{Earnings = \alpha_0 + \alpha_1 Education}$

You find the value of α_0 to be 2000 and the value of α_1 to be 900. How would you interpret these values?

Now imagine you change the measurement of earnings to hundreds of dollars. How would this change your estimates of α_0 and α_1?

What about if you changed the measurement of education to months?

Now, you want to estimate the change in earnings for someone who goes from High School (Education = 0) to Graduate School (Education = 6). What is that change?

🅿️ Nonlinear Terms

Now you run the same regression, but you add a nonlinear term:

$$\boldsymbol{Earnings = \alpha_0 + \alpha_1 Education + \alpha_2 Education^2}$$

Where earnings are measured in dollars and education is measured in years.

You estimate the value of $\boldsymbol{\alpha_0}$ is 2000, $\boldsymbol{\alpha_1}$ is 1000, and $\boldsymbol{\alpha_2}$ is -125.

- What is the marginal effect of education on earnings?
- Does the function have a local minimum or maximum? Where?
- What is the full relationship between education and earnings?
- Now, you want to estimate the change in earnings for someone who goes from High School (Education = 0) to Graduate School (Education = 6). What is that change?

🅿️ Interaction Terms

You run the following regression on how tutoring affects math test scores:

$$\boldsymbol{Score = \beta_0 + \beta_1 Tutoring + \beta_2 LowIncome + \beta_3(Tutoring \times LowIncome)}$$

Where the variable ***LowIncome*** is binary and represents if a student's family's income falls below the poverty line.

You estimate the following:

- $\beta_0 = 0.50$
- $\beta_1 = 0.30$
- $\beta_2 = -0.15$
- $\beta_3 = 0.25$

- How would you interpret each of these terms?
- What is the marginal effect of tutoring on test scores?
- How does the effect of tutoring differ between those considered low-income and those not?
- Sketch a rough graph and label each coefficient.

🅿️ Interaction Terms: Stata Practice!

- Get into groups of 2 or 3 and open stata's automobile dataset (sysuse auto.dta)
- Run the following regression:

$$\boldsymbol{Price = \beta_0 + \beta_1 Weight + \beta_2 Foreign + \beta_3 (Foreign \times Weight)}$$

Hint: you may have to create a new variable!

```
. reg price weight foreign foreignwt

      Source |       SS           df       MS      Number of obs   =        74
-------------+----------------------------------   F(3, 70)        =     26.20
       Model |   335885357         3   111961786   Prob > F        =    0.0000
    Residual |   299180039        70  4274000.55   R-squared       =    0.5289
-------------+----------------------------------   Adj R-squared   =    0.5087
       Total |   635065396        73  8699525.97   Root MSE        =    2067.4

------------------------------------------------------------------------------
       price | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      weight |   2.994814   .4163132     7.19   0.000     2.164503    3.825124
     foreign |  -2171.597   2829.409    -0.77   0.445    -7814.676    3471.482
   foreignwt |   2.367227   1.121973     2.11   0.038      .129522    4.604931
       _cons |  -3861.719   1410.404    -2.74   0.008    -6674.681   -1048.757
------------------------------------------------------------------------------
```

Additional Practice?

- 📝 **Playing with logs**
- 📝 **Interpreting Regression Coefficients**
- 📝 **Worksheet: Regression Mechanics**
- 📝 **Working through magnitudes, quadratic, and BS: Climate change**