Using the t-test - Quiz

Bookmark this page

Finger Exercises due Nov 9, 2020 18:30 EST Past Due

Question 1

0.0/1.0 point (graded)

Consider the hypothesis test $H_0 : R\beta = c$, where $R = I$ is the identity matrix. What does $P\left(H_0\right) = P\left(R\beta = c\right)$ capture?

- The product of the probabilities $\prod_i P\left(beta_i = c_i\right)$
- The probability that $\beta_i = c_i$ holds for some i
- The probability $\beta_i = c_i$ holds simultaneously for all i ✔

Explanation

The matrix equivalency $I\beta = c$ reads row-wise as $\beta_i = c_i$. The hypothesis test H_0 represented by $I\beta = c$ tests that $\beta_i = c_i$ jointly or simultaneously. Note that this is a more stringent test then testing that a particular $\beta_i = c_i$.

Show answer

Question 2

0.0/1.0 point (graded)

True or False: If you are interested in testing a hypothesis of the form $H_0 : \beta_j > c$ vs. $H_1 : \beta_j \leq c$. The F-test and t-test are equivalent, since the t-test static and critical values are the square root of those for the F-test.

- True
- False ✔

Explanation

$H_0 : \beta_j > c$ is a one-sided hypothesis, and therefore you need a t-test. And in this case, the F-test and t-test are not equivalent.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Finger Exercises due Nov 9, 2020 18:30 EST Past Due

Question 1

0.0/1.0 point (graded)

Which of the following hypotheses is the default F-test statistic computed by R testing? (Select all that apply)

- [ ] The hypothesis that $I\beta = \vec{0}$, where β are all the coefficients other than the intercept, and $\vec{0}$ is a vector of 0s ✔
- [ ] The hypothesis that all the regressions coefficients with the exception of the intercept are equal to 0 simultaneously ✔
- [ ] The hypothesis that any of the coefficients on your regressors is equal to 0.
- [ ] The hypothesis that all of your coefficients are equal to 0 simultaneously.

Explanation

The default F-stat reported by R tests the null hypothesis that all your coefficients, with the exception of the intercept are equal to 0. Recall that the set of coefficients on your regressors does not include the intercept, so the hypothesis that all of your coefficients are equal to 0 is false.

Show answer

Submit You have used 0 of 2 attempts

Question 2

0.0/1.0 point (graded)

True or False: If you have only one regressor, the default F-test statistic and t-test statistic reported by R for the coefficient on that regressor will be equivalent. Here "equivalent" means whether you derive the same conclusion of the test. (Intutievely, think whether there is a relationship between the t and the F statistics.)

- True ✔
- False

Explanation

The F-test statistic computed by R is testing whether each of the coefficients on your regressors is equal to 0, whereas the reported t-test statistic is testing that the specific coefficient is equal to 0. Therefore, if you only have one regressor, these will be equivalent. In particular, the reported F-test statistic will be the square of the reported t-test statistic.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Suppose you are interested in the effect of education on wages. To this goal, you run the following regression in R:

$$\log(\text{wage})_i = \alpha + \beta \text{years of education}_i + \epsilon_i$$

where i indexes the individuals in your sample. The reported t-test statistic on β is 2.09, and the reported coefficient is $\hat{\beta} = 0.10$. Furthermore, you know your sample size is large enough for the t-distribtion to be very close to a normal distribution.

What can you conclude given this information?

- You reject the null that $\beta = 0$ at the 95% confidence level. ✔
- You are 95% sure that β is positive.
- You are 95% sure that $\beta = 0$.
- You are 95% sure that $\beta = 0.10$.
- You accept the null that $\beta = 0$ at the 95% confidence level.

Explanation

The reported t-test statistic tests the specific null hypothesis $H_0 : \beta = 0$, against the alternative hypothesis $H_1 : \beta \neq 0$. If you want to test a different hypothesis (ex. $\beta > 0$) you need to construct a different test. In practice, you can use R for both of these. Look up the R command `linearHypothesis` for more details.

Finger Exercises due Nov 9, 2020 18:30 EST **Past Due**

Question 1

0.0/1.0 point (graded)

Let σ_{Neyman} denoted the Neyman standard error, and σ_{OLS} denote the standard error given by your OLS regression. Which of the following statements **are true?** Select all that apply.

- [ ] If your sample is very large, σ_{Neyman} is close to σ_{OLS}. ✔
- [ ] If your sample is very large, $\sigma_{\text{Neyman}} \neq \sigma_{\text{OLS}}$ you should use σ_{OLS}
- [ ] If you have much fewer females than males in your sample, it's still true that $\sigma_{\text{Neyman}} = \sigma_{\text{OLS}}$.
- [ ] If the number of males in your sample is equal to the number of females in your sample, and your sample size is very small, $\sigma_{\text{Neyman}} = \sigma_{\text{OLS}}$. ✔

Explanation

Recall that the Neyman standard errors adjust for the sample of each of the two groups. Whereas, OLS standard errors only take into account the size of the overall sample. Furthermore, the bias disappears if your samples are large enough (so B is not true). However, if you have a very unbalanced sample (i.e much fewer females than males, then $\sigma_{\text{Neyman}} \neq \sigma_{\text{OLS}}$).

Show answer

Finger Exercises due Nov 9, 2020 18:30 EST Past Due

Suppose you have a categorical variable, which denotes the treatment group to which individual *i* has been assigned. You are interested in testing whether providing people with incentives to stay in school will help them stay in school, and also whether providing people with information about the importance of schooling will get them to stay in school longer.

- $T_i = 0$ if individual i has been assigned to the control group.
- $T_i = 1$ if individual i has been assigned to a treatment group that receives incentives for staying in school.
- $T_i = 2$ if individual i has been assigned to a treatment group that receives information on the importance of staying in school.
- $T_i = 3$ if individual i has been assigned to a treatment group that receives both incentives and information

---

Question 1

0.0/1.0 point (graded)

True or False: The problem with including the categorical treatment variable as described above is that it will be collinear and R will refuse to run your regression.

- True
- False ✔

Explanation

Unless you are including a variable that is a scalar multiple of your treatment vector in the regression, there is nothing for your variable to be collinear with. The main problem with including the variable as a categorical treatment variable, like the one described above, is that the coefficient is not interpretable, since the values of the treatment variable themselves are meaningless, i.e. you could have just as easily coded treatment 2 as treatment 3 since the treatment number (category) does not describe anything inherent about the treatment itself. This is why you should encode your categorical variable as dummy variables, as the coefficients on the dummy variables are easily interpretable and allows you to see the impact of each of the treatments you are testing.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

True or False: Suppose you encode your treatment variable T, and generate a set of dummy variables T^j for $j = 0, 1, 2, 3$, where $T^j = 1$ if $T = j$ and 0 otherwise. If you include all 4 indicators as regressors in your model you will have a collinearity problem. (Note: The model includes an intercept.)

- [ ] True ✔
- [ ] False

False

Explanation
Recall the table Prof. Duflo drew in class. One of the variables can be written as a linear combination of the others so you need to omit one of them in order to avoid a collinearity problem.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Question 3

0.0/1.0 point (graded)

Keeping in mind that you are interested in the effect of these different treatments on education, which of the dummy variables would make sense to exclude from your model?

T^0 ✔

- ◯ T^1
- ◯ T^2
- ◯ T^3

Explanation

Suppose you included T^0 from your model, and omitted one of the other indicators. Think about the interpretation of your coefficient for a second. Your coefficients are estimates of the impact of T^j relative to the omitted group. Given that you are interested in the effect of the different treatments on education, it makes sense to compare the treatment groups to the control group, rather than comparing the control group and 2 of the treatment groups to the omitted treatment group.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Finger Exercises due Nov 9, 2020 18:30 EST **Past Due**

Suppose you run the following regression:

$$Y_i = \alpha + \beta M_i + \epsilon_i$$

where Y_i denotes the standardized SAT score of person i and M_i be an indicator equal to 1 if person i belongs to a minority, and 0 otherwise. You have data from a sample of students from your university. You load it into R, and run the regression, and get the following output:

```
Call:
lm(formula = score ~ minority, data = sample)

Residuals:
    Min      1Q  Median      3Q     Max
-3.6700 -0.6754  0.0043  0.6722  4.0304

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.038206   0.007292  142.38   <2e-16 ***
minorityTRUE -0.315742   0.013368  -23.62   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1 on 26784 degrees of freedom
Multiple R-squared:  0.0204,    Adjusted R-squared:  0.02037
F-statistic: 557.9 on 1 and 26784 DF,  p-value: < 2.2e-16
```

Question 1

0.0/1.0 point (graded)

What is the mean standardized SAT score for whites (non-minorities)?

(Please round your answer to the second decimal place, i.e. if your answer is 5.222, round to 5.22 and if it is 5.229, round to 5.23)

Answer: 1.038206

Explanation

The mean standardized SAT score for whites can be found by looking at the estimate for the intercept, since this is the value of y when the dummy variable for minority, M_i is set to 0. From the table, we can see that this is 1.038206, which rounds to 1.04.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

What is the mean SAT score for people who belong to a minority?

(Please round your answer to the second decimal place, i.e. if your answer is 5.222, round to 5.22 and if it is 5.229, round to 5.23)

Answer: 0.722464

Explanation

With a dummy variable, the coefficient represents the difference in means between the two groups. The constant (intercept) is the mean of the outcome variable when $M_i = 0$. In other words, the constant gives the mean SAT score for non-minorities. The difference in means (which you can think of as the intercept shift), is given by $\hat{\beta}$. So the mean SAT score for minorities is the sum of $\hat{\alpha} + \hat{\beta} = 1.038206 - 0.315742 = 0.722464$, which rounds to 0.72

Show answer

Submit

You have used 0 of 2 attempts

Answers are displayed within the problem

Finger Exercises due Nov 9, 2020 18:30 EST **Past Due**

Back to our SAT scores by ethnicity example. You decide you want to do something to help minorities score better on the SAT. To that goal, you decide to run an SAT prep course, and are interested in seeing whether or not that had an impact.

Suppose you run the following regression:

$$Y_i = \alpha + \beta D_i + \gamma M_i + \delta D_i M_i + \epsilon_i$$

where Y_i denotes the SAT score of person i and D_i be an indicator equal to 1 if person i took an SAT prep course, and 0 otherwise, and M_i is an indicator equal to 1 if person i belongs to a non-white minority, and 0 otherwise. You have data from a sample of students from your university. You load it into R, and run the regression, and get the following output:

```
Call:
lm(formula = score ~ minority + presat + presat * minority, data = sample)

Residuals:
    Min      1Q  Median      3Q     Max
-4.2579 -0.6729 -0.0016  0.6800  4.5101

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               0.998960   0.008165 122.339   <2e-16 ***
minorityTRUE             -0.300283   0.014772 -20.328   <2e-16 ***
presatTRUE                0.203320   0.018322  11.097   <2e-16 ***
minorityTRUE:presatTRUE   0.002893   0.033334   0.087    0.931
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9976 on 26782 degrees of freedom
Multiple R-squared:  0.0252,    Adjusted R-squared:  0.0251
F-statistic: 230.8 on 3 and 26782 DF,  p-value: < 2.2e-16
```

For the following questions, please round your answer to the second decimal place (i.e. if your answer is 2.006, round to 2.01 and if it is 2.001, round to 2.00 and if it is 2 express is as 2.00)

Based on these estimates, what is the mean SAT score for white students (non-minorities) who did not take a prep course?

Answer: 0.998960

What is the difference in scores between non-minorities who took the prep course and non-minorities who didn't?

Answer: 0.203320

For those who didn't take the prep course, what is the difference in scores between minorities and non-minorities? *Note: this should be a negative number.*

Answer: -0.300283

What is the difference in the effect of the prep course on SAT scores for minorities relative to non-minorities?

Answer: 0.00283

Explanation

$\hat{\alpha}$ denotes the mean SAT score for the group that has $D_i = M_i = 0$, so $\hat{\alpha}$ the mean SAT score for white students who did not take a prep course. In the regression output, this corresponds to the estimate for (Intercept).

$\hat{\beta}$ is the difference in scores between non-minorities who took the prep course and non-minorities who didn't. In the regression output, this corresponds to the estimate for the coefficient "presatTRUE"

$\hat{\gamma}$ is the difference in scores between minorities and non-minorities who didn't take the prep course. In the regression output, this corresponds to the estimate for the coefficient "minorityTRUE"

$\hat{\delta}$ denotes the additional effect of the prep course for minorities, so $\hat{\delta}$ is the difference in the effect of the prep course on SAT scores between minorities and non-minorities. In the regression output, this corresponds to the estimate for the coefficient "minorityTRUE:presatTRUE"

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Going back to the previous example. For reference, here is the model we considered:

$$Y_i = \alpha + \beta D_i + \gamma M_i + \delta D_i M_i + \epsilon_i$$

where Y_i denotes the SAT score of person i and D_i be an indicator equal to 1 if person i took an SAT prep course, and 0 otherwise, and M_i is an indicator equal to 1 if person i belongs to a non-white minority, and 0 otherwise.

When you run the regression in R, you get the following output:

```
Call:
lm(formula = score ~ minority + presat + presat * minority, data = sample)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.2579 -0.6729 -0.0016  0.6800  4.5101 

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)    
(Intercept)              0.998960   0.008165 122.339   <2e-16 ***
minorityTRUE            -0.300283   0.014772 -20.328   <2e-16 ***
presatTRUE               0.203320   0.018322  11.097   <2e-16 ***
minorityTRUE:presatTRUE  0.002893   0.033334   0.087    0.931    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9976 on 26782 degrees of freedom
Multiple R-squared:  0.0252,	Adjusted R-squared:  0.0251 
F-statistic: 230.8 on 3 and 26782 DF,  p-value: < 2.2e-16
```

Question 1

0.0/1.0 point (graded)

Given your estimates $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$, which of the following expression gives your estimate of the mean SAT score for non-minorities who took the prep course?

- [ ] $\hat{\alpha} + \hat{\delta}$
- [ ] $\hat{\alpha} + \hat{\beta}$ ✔
- [ ] $\hat{\alpha} + \hat{\gamma}$
- [ ] $\hat{\alpha} - \hat{\beta}$

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

Going back to the empirical output, what is your empirical estimate of the mean SAT score for non-minorities who took the prep course? Please round your answer to the second decimal place (7.8654 is equivalent to 7.87).

Answer: 1.20228

Explanation

As you found out in the last segment, $\hat{\alpha} = 0.998960$ and $\hat{\beta} = 0.203320$. So, based on question 1, $\hat{\alpha} + \hat{\beta} = 1.20228$.

Show answer

Submit You have used 0 of 2 attempts

Finger Exercises due Nov 9, 2020 18:30 EST Past Due

Going back to the previous example regarding SAT scores. For reference, here is the model we are considering:

$$Y_i = \alpha + \beta D_i + \gamma M_i + \delta D_i M_i + \epsilon_i$$

where Y_i denotes the SAT score of person i and D_i be an indicator equal to 1 if person i took an SAT prep course, and 0 otherwise, and M_i is an indicator equal to 1 if person i belongs to a non-white minority, and 0 otherwise.

Question 1

0.0/1.0 point (graded)

Given your estimates $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$, which of the following expression gives your estimate of the mean SAT score for minorities who took the prep course?

- ◯ $\hat{\alpha} + \hat{\beta} + \hat{\gamma} + \hat{\delta}$ ✔
- ◯ $\hat{\alpha} + \hat{\beta} + \hat{\gamma} - \hat{\delta}$
- ◯ $\hat{\alpha} + \hat{\gamma} + \hat{\delta}$
- ◯ $\hat{\alpha} - \hat{\beta} + \hat{\gamma}$

Hint: Draw the 3x3 by table Prof. Duflo explained in class to work this out

No Prep	Prep	
$\hat{Y}_{00} = \hat{\alpha}$	$\hat{Y}_{01}$	Non-Minority
$\hat{Y}_{10}$	$\hat{Y}_{11}$	Minority

$$\hat{\beta} = \bar{Y}_{01} - \bar{Y}_{00}$$
$$\hat{\gamma} = \hat{Y}_{10} - \hat{Y}_{00}$$
$$\hat{\delta} = (\hat{Y}_{11} - \hat{Y}_{01}) - (\hat{Y}_{10} - \hat{Y}_{00})$$

We want to find $\bar{Y}_{11}$, since:

$$\hat{\delta} = (\bar{Y}_{11} - \bar{Y}_{10}) - (\bar{Y}_{01} - \bar{Y}_{00}) = (\bar{Y}_{11} - \bar{Y}_{10}) - \hat{\beta}, [mathjax] wehavethat [mathjax] \bar{Y}_{11} = \hat{\delta} + \hat{\beta} + \bar{Y}_{10}$$

And since

$$\hat{\gamma} = \bar{Y}_{10} - \bar{Y}_{00} = \bar{Y}_{10} - \hat{\alpha}$$

we have that

$$\bar{Y}_{10} = \hat{\gamma} + \hat{\alpha}$$

Finally, substituting we get the expression

$$\hat{\alpha} + \hat{\beta} + \hat{\gamma} + \hat{\delta}$$

Finger Exercises due Nov 9, 2020 18:30 EST **Past Due**

Question 1

0.0/1.0 point (graded)

In Card's Mariel boatlift study, his outcome of interest was:

- ◯ the labor market outcomes for people who already lived in Miami prior to the boatlift ✔
- ◯ the percentage of people from Cuba who arrived in Miami and stayed in Miami
- ◯ the labor market outcomes for the Cubans who arrived in Miami after the boatlift
- ◯ the number of people who arrived in Miami from Cuba

Explanation

As stated in the lecture, Card was interested in understanding whether the influx of immigrants from Cuba affected the labor market outcomes of people who lived in Miami anyway.

Show answer

Submit You have used 0 of 2 attempts

Question 2

0.0/1.0 point (graded)

True or False: Similar to the Card and Krueger study on minimum wage, which compares New Jersey and Pennsylvania, Card compares Miami to **one** other city of similar size.

- True
- False ✔

Explanation

This statement is false. For the Mariel boat study, Card compares Miami to a bunch of cities that are reasonably similar to Miami in terms of size, education, racial makeup, et cetera. Later in the lecture, Prof. Duflo discusses the merits of using a weighted average of other cities such that the weighted average looks very similar to Miami; however, this was not the methodology used in Card's original study.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

Finger Exercises due Nov 9, 2020 18:30 EST **Past Due**

Question 1

0.0/1.0 point (graded)

Which of the following is **not** a finding of Duflo (2001)?

- ◯ The growth in education levels was higher in places where more schools were built.
- ◯ Younger people were on average more educated than older people.
- ◯ Prior to the building of the schools, the regions where schools were built were similar to the regions where the school's weren't built ✔
- ◯ In regions where a lot of schools were built, people were on average less educated prior to the construction of the schools.
- ◯ None of the above

Explanation

The results in Professor Duflo's table suggest that schools were built in regions where people were on average less educated and where younger people were on average more educated than older people. Finally, the main result of Duflo (2001) is that the growth in education levels was higher in places where more schools were built.

Question 2

0.0/1.0 point (graded)

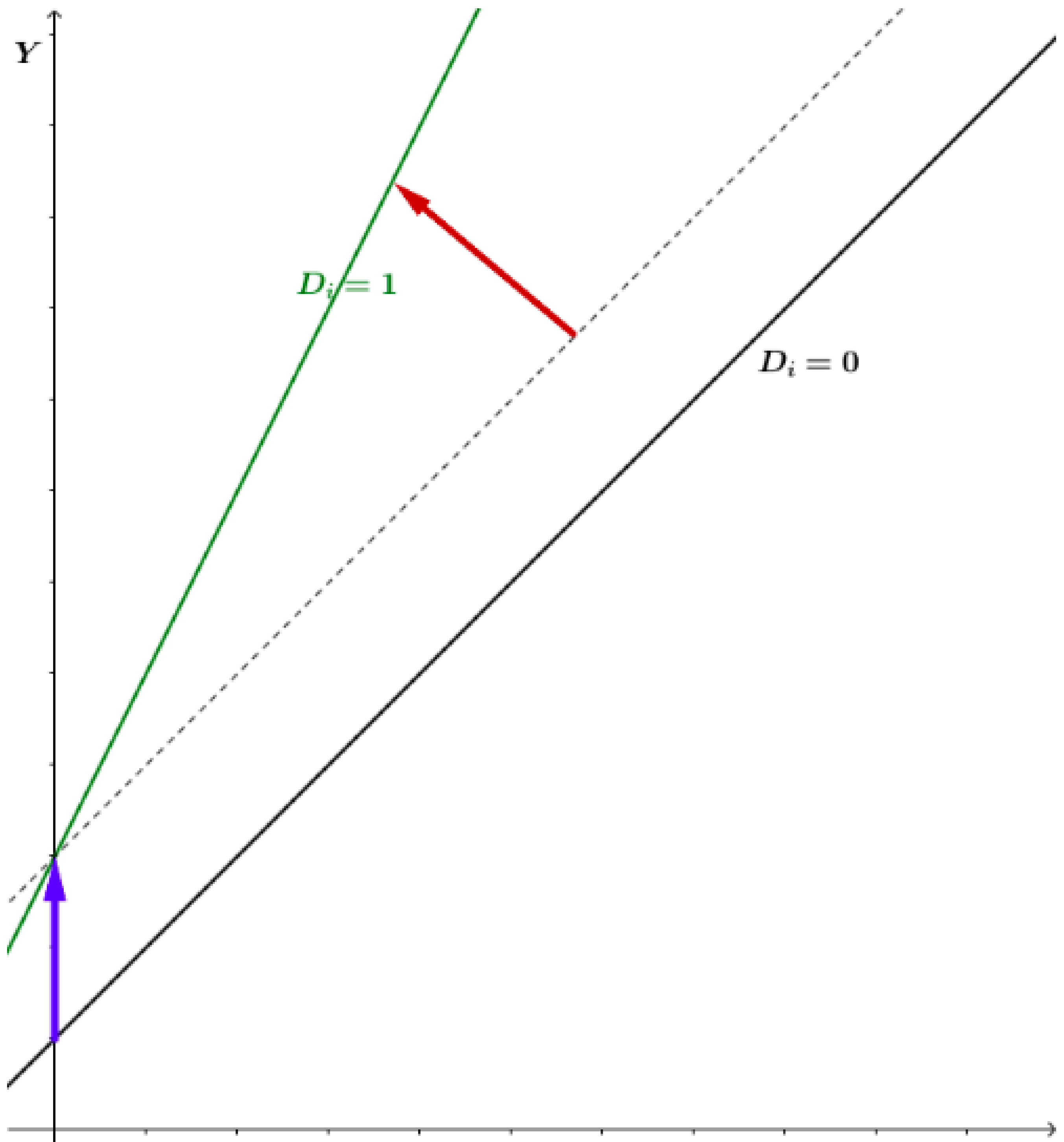
What is the underlying assumption in Duflo (2001)?

- ◯ In the absence of the program, there would not have been any growth in education levels.
- ◯ In the absence of the program, the difference in education levels between regions where many schools were built and regions where fewer schools were built would have been constant over time. ✔
- ◯ In the absence of the program, regions where a lot of schools were built would've been comparable to regions where few schools were built.
- ◯ In the absence of the program, there would not have been a change in education levels over time.
- ◯ None of the above

Explanation

The underlying assumption of Duflo (2001) is that in the absence of the program, the difference in education levels between regions where many schools were built and regions where fewer schools were built would have been constant over time. This is often referred as the parallel trends assumption, or the common trends assumption. More generally, the underlying assumption in difference-in-differences designs is that in the absence of the program, the difference in outcomes between places that received the program and places that didn't would've stayed constant, (i.e given their underlying differences, they follow parallel trends).

Suppose you are estimating a model of the form $Y_i = \alpha + \beta D_i + \gamma X_i + \delta D_i X_i + \epsilon_i$, where X is a continuous regressor, and D is a dummy variable. The figure below provides a graphical representation of this model using the same framework Prof. Duflo used in class.



0.0/1.0 point (graded)

For each of the following, determine which parameter corresponds to the respective feature of the above graph?

I. What parameter is represented by the blue arrow? *(write a for α, b for β, g for γ, and d for δ.)*

Answer: b

II. What is the slope of the solid green line (in terms of the parameters in the model)?

- $\gamma + \delta$ ✔
- $\beta + \delta$
- $\gamma + \alpha$
- $\alpha + \beta$

Explanation

As Prof. Duflo illustrated in class, the intercept of the solid black line is given by α, the change in the intercept between $D_i = 0$ and $D_i = 1$ is given by the coefficient on D_i, β. The slope of the black solid line, is given by γ the coefficient on X_i. The coefficient on interaction term tells us the extent to which D_i changes the slope of the outcomes as a function of your continous regressor, so the red arrow represents δ. So the solid green line has slope $\gamma + \delta$, and intercept $\alpha + \beta$.

Question 1

0.0/1.0 point (graded)

Why does Prof. Duflo include year-of-birth fixed effects in her fixed effects specification?

- ◯ To control for inherent differences in education levels across time. ✔
- ◯ To account for differences in the age distribution across regions.
- ◯ To look at the impact of year-of-birth on education levels.
- ◯ None of the above

Explanation

Year of birth fixed effects are a set of dummy variables indicating the year-of-birth (one dummy for each year-of-birth (yob)). The coefficient on a given yob dummy is just the mean education (in years) for children born in that year. She includes these yob dummies to separate the effect of school building from the effect of being young. Since the young are more educated anyway, if she compared the old to the young without controlling for their yob, her estimator would conflate the actual effect of school building on education with the trend of increasing education levels across time.

Show answer

Question 2

0.0/1.0 point (graded)

Let's continue with the INPRES example. Recall the following model:

$$S_{ijk} = c_1 + \alpha_{1j} + \beta_{1k} + (P_j * T_i)\,\gamma + \epsilon_{ijk}$$

What is the interpretation of the coefficient γ in this model?

- ◯ γ denotes the impact of year of birth on the effect of each additional school built.
- ◯ γ denotes the difference in means between regions with a lot of schools and regions with fewer schools.
- ◯ γ denotes the difference in means between regions with a higher density of young people relative to old people.
- ◯ γ denotes the impact of one more school built in the region on the difference in years of education between the young and the old. ✔

Explanation
γ is the coefficient on the interaction term between the dummy for young and the continuous measure of the number of schools built per 1000 children. Therefore, γ is the additional effect of each school for the young relative to the old.

Question 1

0.0/1.0 point (graded)

Which of the following statements are true? Select all that apply.

- [ ] β_1 is the elasticity of wage with respect to education.
- [ ] Each additional year of education leads to a $(\beta_1 * 100)\,\%$ change in wages. ✔
- [ ] β_2 is the elasticity of wage with respect to mother's education. ✔
- [ ] A 1% change in education leads to a $\beta_1\%$ change in wages.
- [ ] A 1% change in mother's education leads to a $\beta_2\%$ change in wages. ✔

Explanation

When your outcome is in logs and your regressor is in logs, the coefficients represent elasticities: your coefficients measure the $\%$ change in your outcome as a result of a 1% **change in your regressor.** If your outcome is in logs, but your regressor is not, the coefficient represents the $\%$ in your outcome resulting from a **unit increase in your regressor.** So in this example, since the model includes S_i and $\log(P_i)$, each additional year of education leads to a $(\beta_1 * 100)\,\%$ change in wages and a 1% change in mother's education leads to a $\beta_2\%$ change in wages. β_2 is also the elasticity of wage with respect to mother's education.

Question 1

0.0/1.0 point (graded)

Which of the following statements are true? Select all that apply.

- [ ] β_1 is the elasticity of wage with respect to education.
- [ ] Each additional year of education leads to a $(\beta_1 * 100)\,\%$ change in wages. ✔
- [ ] β_2 is the elasticity of wage with respect to mother's education. ✔
- [ ] A 1% change in education leads to a $\beta_1\%$ change in wages.
- [ ] A 1% change in mother's education leads to a $\beta_2\%$ change in wages. ✔

Explanation

When your outcome is in logs and your regressor is in logs, the coefficients represent elasticities: your coefficients measure the $\%$ change in your outcome as a result of a 1% **change in your regressor.** If your outcome is in logs, but your regressor is not, the coefficient represents the $\%$ in your outcome resulting from a **unit increase in your regressor.** So in this example, since the model includes S_i and $\log(P_i)$, each additional year of education leads to a $(\beta_1 * 100)\,\%$ change in wages and a 1% change in mother's education leads to a $\beta_2\%$ change in wages. β_2 is also the elasticity of wage with respect to mother's education.