



Bookmarks

- [Module 1: The Basics of R and Introduction to the Course](#)
- [Entrance Survey](#)
- [Module 2: Fundamentals of Probability, Random Variables, Distributions, and Joint Distributions](#)
- [Module 3: Gathering and Collecting Data, Ethics, and Kernel Density Estimates](#)
- [Module 4: Joint, Marginal, and Conditional Distributions & Functions of Random Variable](#)

Module 10: Practical Issues in Running Regressions, and Omitted Variable Bias > Practical Issues in Running Regressions > Regressions with Categorical Variables - Quiz

Regressions with Categorical Variables - Quiz

🔖 Bookmark this page

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

1/1 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.

- ▶ [Module 5: Moments of a Random Variable, Applications to Auctions, & Intro to Regression](#)
- ▶ [Module 6: Special Distributions, the Sample Mean, the Central Limit Theorem, and Estimation](#)
- ▶ [Module 7: Assessing and Deriving Estimators - Confidence Intervals, and Hypothesis Testing](#)
- ▶ [Module 8: Causality, Analyzing Randomized Experiments, & Nonparametric Regression](#)
- ▶ [Module 9: Single and Multivariate Linear Models](#)
- ▼ [Module 10: Practical Issues in Running](#)

☐ a. True

☒ b. 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. As Prof. Duflo discussed in class, 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.

Submit

You have used 1 of 1 attempt

✓ Correct (1/1 point)

Question 2

1/1 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.

Regressions, and Omitted Variable Bias

Practical Issues in Running Regressions

due Dec 5, 2016 05:00 IST



Omitted Variable Bias

due Dec 5, 2016 05:00 IST



Module 10: Homework

due Nov 28, 2016 05:00 IST



☒ a. True ✓

☐ b. 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.

Submit

You have used 1 of 1 attempt

✓ Correct (1/1 point)

Question 3

1/1 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?

☒ a. T^0 ✓

☐ b. T^1

☐ c. T^2 ☐ d. 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.

[Submit](#)

You have used 1 of 2 attempts

✓ Correct (1/1 point)

Discussion**Topic:** Module 10 / Regressions with Categorical Variables - Quiz[Show Discussion](#)



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

