

Lista 6

Do-File and Answer Key are due October 5, 2017.

Exercise 1. Please undertake the following routine in Stata. What do you observe from the results of both regressions for the coefficient estimates? How are these results related? Please explain.

1. set obs 100
2. gen x = invnorm(uniform())
3. gen e1 = invnorm(uniform())
4. gen y1 = 2 + 3*x + e1
5. reg y1 x
6. gen e2 = invnorm(uniform())
7. gen y2 = 2 + 3*x + e2
8. reg y2 x

Exercise 2. Now let us assume that the explanatory variable, X, is measured with error. This measurement error is relatively small (.01). What are the consequences for the regression results reported in Exercise 1?

Exercise 3. Now let us assume that the dependent variable, Y, is measured with error. This measurement error is relatively small (.01). What are the consequences for the regression results reported in Exercise 1?

Exercise 4. Based on your analysis of the results in Exercises 2 and 3, what do you conclude? How do your conclusions related to the regression assumptions outlined by Gujarati?

Parte II. The Multiple Regression Model

Exercise 5. Now let's add a second explanatory variable, Z. One of the assumptions of the regression model is that there is no exact collinearity between the explanatory variables in a multiple regression model. Please create a variable Z that does not violate this assumption. Now, do a scatterplot to analyze the relationship between X and Z.

Exercise 6. Please re-do the analysis reported in Exercise 1 adding a second explanatory variable to the model. What do you observe from the results of both regressions for the coefficient estimates? How is the interpretation of the coefficients different from Exercise 1?

Exercise 7. Now let's add a third explanatory variable, V. Please create a variable V that is a dummy variable where precisely 40% of the observations are coded as 1 and 60% are coded as 0. Please re-do the analysis reported in Exercise 1 adding V as a second explanatory variable to the model. What do you observe from the results of both regressions for the coefficient estimates? How is the interpretation of the coefficients different from Exercise 1?