

Elements of Econometrics. 2022-2023.
Class 8. Dummy Variables

Problem 1. Discuss how dummy variables can be used to test

- (a) change in intercept;
- (b) change in slope;
- (c) changes in both intercept and slope.

Problem 2. A student decided to investigate the market of private mathematics teachers in Moscow, with particular interest to those who can teach in English. He took a random sample of 30 profiles of teachers who provide private teaching in math (taken from population of 300 profiles registered in certain internet site) and run some regressions trying to find factors influencing the prices of teaching ($PRICE_i$ - price of a standard two-hour lesson in thousands of roubles, $DIST_i$ - distance in the number of metro stations from the center of Moscow to the teacher's place, $HOME_i$ - dummy variable indicating visit of the tutor to the client, ENG_i - dummy variable indicating ability to teach the subject in English):

$$\widehat{PRICE}_i = 6.59 - 0.16DIST_i \quad R^2 = 0.185$$

(0.49) (0.06) (1)

$$\widehat{PRICE}_i = 4.51 + 2.54HOME_i \quad R^2 = 0.40$$

(0.40) (0.58) (2)

$$\widehat{PRICE}_i = 5.13 - 0.08DIST_i + 1.95HOME_i + 0.07DIST * HOME_i \quad R^2 = 0.437$$

(0.64) (0.06) (0.95) (0.07) (3)

$$\widehat{PRICE}_i = 4.52 - 0.08DIST_i + 2.18HOME_i + 1.58ENG_i - 0.39HOME * ENG_i \quad R^2 = 0.553$$

(0.61) (0.06) (0.75) (0.76) (1.09) (4)

(a) What is the meaning of equations (1) and (2)? What are advantages and disadvantages of these equations? What is the difference in the meaning and in the assumptions (explicit or implicit) used for equation (3) as to compare to equation (1)? Give interpretation to all coefficients of equation (4).

(b) Is equation (3) significant? Is factor "distance" (the variables $DIST_i$ and $DIST * HOME_i$ taken together) significant in equation (3)? Is factor "teaching at student's place" (the variables $HOME_i$ and $DIST * HOME_i$ taken together) significant in equation (3)? Are all dummy variables taken together in equation (4) significant?

(c) Another approach to the analysis of influence of qualitative factors is Chow test. Explain how you will do Chow test for the analysis of the influence of the place of teaching. Can you predict the result of this test? Explain how you will do Chow test for the analysis of the influence of both the place of teaching and ability to teach in English. Write the equation with dummy variables $H = HOME$ and $E = ENG$ that allows to do F-test for restrictions that is equivalent to the last of the two Chow tests.

Problem 3 A student of ICEF preparing her diploma paper studies the dependence of the prices on the paintings P_i (in thousands of dollars) on various factors, in particular on the "age" of the painting AGE_i (in decades from the current year to the year of creation of the piece of art) and on the size of the canvas S_i (in square feet). She collected data on 19 paintings sold at a certain auction to estimate equation (1)

$$\hat{P}_i = 2.35 + 0.028 \cdot AGE_i + 0.037 \cdot S_i \quad R^2 = 0.325$$

(0.57) (0.011) (0.020) $RSS = 0.417$ (1)

Correlation between variables AGE_i and S_i was -0.93 .

(a) Give the interpretation to all coefficients of this equation.

(b) The supervisor advised the student to take into account also the availability of the provenance (a history of possession of an artwork, confirming the authenticity of the object): the variable $PV_i = 1$ if the provenance is present, and $PV_i = 0$ otherwise. The student estimates two additional equations (3), (4)

$$\hat{P}_i = 0.87 + 0.062 \cdot AGE_i + 0.082 \cdot S + 0.36 \cdot PV_i \quad R^2 = 0.554$$

(0.71) (0.016) (0.024) (0.13) $RSS = 0.276$ (2)

and

$$\hat{P}_i = -0.097 + 0.085AGE_i + 0.11 \cdot S + 2.69PV_i - 0.07PV_i \cdot AGE_i - 0.08PV_i \cdot S_i \quad R^2 = 0.691$$

(0.75) (0.017) (0.025) (3.55) (0.05) (0.14) $RSS = 0.191$ (3)

What is the difference in interpretation of coefficients of the equations (1), (2), (3)? Is the availability of the provenance significant according (2) and (3)?

(c) The alternative approach to the detection of the significance of the provenance is Chow test. Describe in details this approach. Suppose that equation (1) estimated for paintings without provenance gives the value of RSS equal to 0.191, while for those with provenance the value of RSS is 0.000122. Do Chow test and compare the results with the tests above in (d).