DXE_EMTR 2023
First assignment (20% of total grade)

Please submit the assignment by 30.10.2023 in the IS MUNI system. You are allowed to work in groups of **maximum size 3**.

For students who were not exposed to programming before or who wish to increase their familiarity with software `R`, attending `DXE_EREC` course is highly recommended.

It is well known that adding a new regressor into a linear regression model may completely change the values of other estimated regression coefficients.

Consider the following model (M1):

$$y_i = \beta_0 + \beta_1 x_{i1} + \epsilon_i$$

and denote the estimated regression coefficients obtained by the ordinary least squares (OLS) as $\hat{\beta}^{(M1)} = (\hat{\beta}_0^{(M1)}, \hat{\beta}_1^{(M1)})$.

Now we add an additional regressor x_2 and the new model (M2) is now the following:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

and denote the estimated regression coefficients obtained by OLS as $\hat{\beta}^{(M2)} = (\hat{\beta}_0^{(M2)}, \hat{\beta}_1^{(M2)}, \hat{\beta}_2^{(M2)})$.

Regression basics

(1) The estimated coefficients $\hat{\beta}_1^{(M1)}$ and $\hat{\beta}_1^{(M2)}$ are in general different. Write a simple code in R to demonstrate that they are indeed different, even asymptotically (with a large sample size).

(2) If the *true correlation* between x_1 and x_2 is exactly zero, however, then the difference between $\hat{\beta}_1^{(M1)}$ and $\hat{\beta}_1^{(M2)}$ tends to go to zero with a large sample size. Demonstrate this with an example in `R`.

(3) If the *sample correlation* between x_1 and x_2 is exactly zero, then the $\hat{\beta}_1^{(M1)}$ and $\hat{\beta}_1^{(M2)}$ are identical. Demonstrate this with an example in `R` and for different sample sizes.[1]

(4) You have illustrated the points (1), (2), (3) on a computer. Now show the points (1), (2) and (3) analytically (that is: with a pencil and a paper).

(5) The two models (M1) and (M2) produce a different fit. What is the difference between R^2_{M1} from the first model and R^2_{M2} from the second model if the true correlation between x_1 and x_2 is close to

(a) 0,
(b) -1,
(c) 1.

Explain your reasoning.

(6) What is the impact of the correlation between x_1 and x_2 on the standard error of ~~$\hat{\beta}_1^{(M1)}$~~ $\hat{\beta}_1^{(M2)}$. Compare with the theoretical predictions.

(7) In light of (6) explain why experimentalists sometimes prefer *an orthogonal design* - if they can choose x_1 and x_2 (e.g. in which points they measure y) they make x_1 and x_2 uncorrelated.

[1] To generate random vectors so that the *sample correlation* is exactly zero may be tricky. One solution is to change the last number in vector of generated realisations of x_2 with the help of root-finding function `uniroot()`. You may, however, come up with your own method.

Maximum likelihood

It is known that the OLS estimator is the Maximum likelihood estimator under homoscedastic normally distributed iid errors.

(1) Write down the log-likelihood function of MLE estimator of the vector $\beta^{(M1)}$.

(2) Write an R-script that optimize this function (you may use `optim()` function). Compare the results to the OLS estimator you get from the `lm()` function.

(3) Run many simulations and show a simulated distribution of $\hat{\beta}_1^{(M1)}$ for different sample sizes (e.g. 100, 1'000, 100'000). Compare the shape of the distribution with the theoretical predictions.

(4) Explore the sensitivity of the results to

   (a) Number of simulations,

   (b) Correlation between x_1 and x_2, (for (b) you may generate data samples according to model (M2))

   (c) violations of the assumption of the homoscedasticity of errors.

Bootstrap

Now generate a sample of size $n = 30$ for which the second model (M2)

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

is true. On this dataset

(1) Use non-parametric bootstrap to get a standard error of $\hat{\beta}_1^{(M1)}$.

(2) Use non-parametric bootstrap to get a standard error of $\hat{\beta}_1^{(M2)}$.

(3) Compare the differences between (1) and (2) and explain.

(4) Read the short paper *Kennedy, P. E. (2001). Bootstrapping student understanding of what is going on in econometrics. The Journal of Economic Education, 32(2), 110-123* and write a short (yes, short) summary on how the bootstrap may be useful for understanding the concept of a sampling variation.

Submit your own work. Make sure that your code runs without errors. Make sure to comment your code and make your best effort to adhere to some reasonable coding standards. Your code must be easy to read. Present your results in a coherent way and whenever possible make use of visualization.