Problem Set 2

Problem Set 2

The model used in the simulation has the following structure:

$$x_{i1} = 1, x_{i2}, x_{i3} \sim \mathcal{N}(0, \Sigma)$$

$$\Sigma = \begin{bmatrix} 5 & 2.5 \\ 2.5 & 4 \end{bmatrix}$$

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i$$

Assume throughout that $n = 1000$. The aim of this problem set is to implement the F-test for joint hypothesis testing and the t-test for the individual coefficients (you can just pick one) and to compare the performance of these tests if the model assumptions upon which the test statistics are derived, are violated.

Implementing the F-statistic and the t-statistic

Assume initially that $\boldsymbol{\beta} = \begin{pmatrix} 1 & 0.1 & 0.1 \end{pmatrix}$ and that $\varepsilon_i \sim \mathcal{N}(0, 10)$.

- Implementing Statistics

Calculate the OLS estimator $\hat{\boldsymbol{\beta}}$, implement the F-Statistic, the t-statistic and compare the values of the F-Statistic and the t-statistic with the relevant quantiles from the distributions. Do the individual Nulls hold/does the global Null hold?

- Coverage Probabilities

Calculate the confidence interval for the individual coefficients (except the constant) and calculate the coverage probabilities.

Simulation Study

Write the code above into a function that returns a logical expression whether the Null was rejected for both the t and the F statistic. Repeat this calculation 1000 times and record your results in a vector. How can we interpret the results?

Simulation study: Violating model assumptions

Re-write the data generating process above to include the option for simulating heteroskedasticity. Come up with a suitable way to simulate heteroskedasticity in the model and run the simulation above again. What is your expectation?