Exercises: Week 4

Econometrics Prof. Conlon

Ulrich Atz

2021-02-25

Problem 1 (Coding Exercise)

This exercise asks you to implement and assess the performance of the bootstrap for the linear regression model. Suppose you have the linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

where,

- $x_i \sim U[0, 2]$
- $\epsilon_i | x_i \sim U[-1, 1]$
- $\beta_0 = \beta_1 = 1$

We ask you to answer the following questions:

a. Write a code that generates i.i.d. samples of sizes $n = 10, 50, 200$ from that distribution, computes (1) the least squares estimator for β, (2) the t-ratio for the least squares coefficient β_1, $t_n = \frac{\hat{\beta}_{1,LS} - 1}{s\hat{.}e.(\hat{\beta}_{1,LS})}$, and (3) the least square residuals $\hat{\epsilon}_i = y_i - \hat{\beta}_{0,LS} - \hat{\beta}_{1,LS} x_i$

b. Write a code for drawing n times at random from the discrete uniform distribution over the estimated residuals $\hat{\epsilon}_1, ..., \hat{\epsilon}_n$ (i.e. with replacement).

c. Use your code from parts (a) and (b) to implement the residual bootstrap - assuming that ϵ_i and x_i are independent - to estimate the 95th percentiles of the respective distributions of $\hat{\beta}_{1,LS}$ and t_n

d. Repeat part (a) for sample size $n = 10, 50, 200$ with 200 replications, where you keep the initial draws of $x_1, .., x_n$ from part (a) and only generate new residuals from their conditional distribution. Compute $\hat{\beta}_{1,LS}$ and the statistic t_n using 200 independent samples of size n. Use your results to compute a simulated estimate for the 95th percentiles of the respective sampling distributions for $\hat{\beta}_{1,LS}$ and t_n.

e. Compare your results from (c) and (d). What do you conclude about the performance of the bootstrap? How does it compare to the 95th percentile of the asymptotic distribution of t_n?

f. See if you can construct a subsampled confidence interval for $n = 200$ with $a_n = 25$. How does it compare to the bootstrapped CI?