Advanced Econometrics II: 2024-2025

Assignment on Bootstrap

Deadline: Friday, January 10 at 17:00, 2025

1 Assignment

The purpose of this assignment is to help you gain an understanding of when the bootstrap method is applicable. Please submit your work on Canvas before **Friday, 17:00 hours**. The submission should include:

- A report that briefly answers the nine questions stated in Section 3 (max. four pages). Use graphs and tables as requested.
- You can work in teams of 3 students. Clearly state your names and student numbers! After the names, include a declaration of originality that "*1. These solutions are solely our own work. 2. We have not made (part of) these solutions available to any other student. 3. We shall not engage in any other activities that will dishonestly improve our results or dishonestly improve or hurt the results of others*".

Other important details:

- You are free to use your preferred programming language (R, Python, Julia, etc.), but you must also upload your code (**do not provide answers as comments in the code**).
- Not all details are given, so ensure you specify and justify your choices (there is more than one correct answer).

2 Introduction

The bootstrap is usually applied to approximate the (finite-sample) distribution of an (asymptotically normal) estimator or test statistic. In this assignment, you investigate how the bootstrap can be used to deliver a diagnostic (misspecification) test. Classical asymptotic theory for (extremum) estimators is based on the asymptotic behavior of the score (S_n) and the information (I_n). For instance, if

$$n^{-1/2} S_n \xrightarrow{d} N(0, J),$$
$$n^{-1} I_n \xrightarrow{p} -H,$$

as the number of observations $n \to \infty$ and particular regularity assumptions hold, we have

$$T_n = \sqrt{n}(\hat{\theta}_n - \theta_0) \approx (n^{-1} I_n)^{-1} n^{-1/2} S_n \xrightarrow{d} N(0, V),$$

with $V = H^{-1}JH^{-1}$. Such assumptions are usually related to (i) the existence of moments (ii) stationarity (iii) non-singular information and (iv) true parameter in the interior of the parameter space. When these assumptions break down, asymptotic normality is no longer guaranteed in general and inference based on Gaussian limit theory is usually invalid

$$T_n = \sqrt{n}(\hat{\theta}_n - \theta_0) \overset{d}{\nrightarrow} N(0, V).$$

Can we exploit the non-Gaussianity of a bootstrap statistic T_n to detect potential failures of the assumptions behind asymptotic normality? The main idea is:

Step 1 Evaluate the distance of the (conditional) distribution of T_n from the limiting Gaussian distribution

Step 2 If such distance is significant, we reject the validity of the regularity assumptions.

Implementation of this idea using the bootstrap:

1. Generate a random sample of m i.i.d. realizations of T_n^*, say $T_{n:1}^*, ..., T_{n:m}^*$;
2. Compute a distance (e.g., Lilliefors, Shapiro Wilk) between the EDF (empirical distribution function) of the T_n^*'s and the Gaussian CDF;
3. Test if such distance is significant using proper critical values.

This idea was also introduced in Beran (1997) [Diagnosing bootstrap success. Annals of the Institute of Statistical Mathematics, 49, 1-24]. Under certain conditions on m and n, it can be shown that standard critical values for the distance can be employed.

2.1 Example 1: Parameter on/near the Boundary

Consider a simple location model with i.i.d. data:

$$y_i = \theta + \varepsilon_i, \qquad \text{with } E[\varepsilon_i] = 0 \text{ and } E[\varepsilon_i^2] = 1,$$

with $\theta \in \Theta = [0, \infty)$. The QMLE is given by

$$\hat{\theta}_n = \max(0, \bar{y}_n). \tag{1}$$

Consider the parametric residual bootstrap

$$y_i^* = \hat{\theta}_n + \varepsilon_i^*, \qquad \varepsilon_i^* \sim N(0, 1),$$

where

$$\hat{\theta}_n^* = \max(0, \bar{y}_n^*).$$

When the true mean is zero, Andrews (2000) [Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. Econometrica, 399-405] showed that the bootstrap is inconsistent, i.e. not asymptotically correct to the first order. This is true for the nonparametric bootstrap based on the empirical distribution function, as well as the parametric bootstrap based on the restricted or unrestricted maximum likelihood estimator.

2.2 Example 2: Infinite Variance

Consider again the location model with i.i.d. data but now with infinite variance:

$$y_i = \theta + \varepsilon_i, \qquad \text{with } E[\varepsilon_i] = 0 \text{ and } E[\varepsilon_i^2] = +\infty,$$

and $\theta \in \mathbb{R}$. Let $\hat{\theta}_n$ denote the sample mean. Consider the bootstrap DGP:

$$y_i^* = \hat{\theta}_n + \varepsilon_i^*, \qquad \varepsilon_i^* \text{ i.i.d. from } \hat{\varepsilon}_i = y_i - \hat{\theta}_n.$$

Knight (1989) [On the bootstrap of the sample mean in the infinite variance case. The Annals of Statistics, 1168-1175] proved that the bootstrap is invalid and that the asymptotic bootstrap distribution of $T_n^* = a_n^{-1} n(\hat{\theta}_n^* - \hat{\theta}_n)$ is non-Gaussian.

2.3 Example 3: Weak Instruments

The DGP is a simple regression with one endogenous regressor:

$$\begin{aligned} y_i &= \beta x_i + \varepsilon_i, \\ x_i &= \pi z_i + u_i, \end{aligned}$$

with $\mathbb{C}(z_i, u_i) = \mathbb{C}(z_i, \varepsilon_i) = 0$, $\mathbb{C}(\varepsilon_i, u_i) \neq 0$ and $\pi \neq 0$ (strong instrument). For simplicity, suppose $\varepsilon_i \sim N(0,1)$, $u_i \sim N(0,1)$ and $z_i \sim N(0,1)$. The IV estimator is given by $\hat{\beta}_n = S_{zx}^{-1} S_{zy}$ with $S_{ab} = n^{-1} \sum_{i=1}^{n} a_i b_i'$ and

$$T_n = \sqrt{n}(\hat{\beta}_n - \beta) = S_{zx}^{-1} \sqrt{n} S_{z\varepsilon} \xrightarrow{d} N(0, \omega^2), \qquad \omega = \frac{1}{\pi} \frac{\sigma_\varepsilon}{\sigma_z}. \tag{2}$$

Consider the (infeasible) residual bootstrap:

$$\begin{aligned} y_i^* &= \hat{\beta}_n x_i^* + \varepsilon_i^*, \\ x_i^* &= \hat{\pi}_{ols} z_i + u_t^*, \end{aligned}$$

with

$$(\varepsilon_i^*, u_i^*) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right).$$

This is considered infeasible because ρ is unknown in practice. For the bootstrapped root, we get

$$T_n^* = \sqrt{n}(\hat{\beta}_n^* - \hat{\beta}_n) = S_{zx^*}^{-1} \sqrt{n} S_{z\varepsilon^*} \xrightarrow{d^*} N(0, \omega^2).$$

However, in the case of a weak instrument, i.e. $\pi = \pi_n = cn^{-1/2}$, it can be shown that both T_n and T_n^* are both non-Gaussian and that bootstrap inference based on T_n^* is invalid.

2.4 The Bootstrap Test

Suppose

$$T_n^* = \frac{\hat{\theta}_n^* - \hat{\theta}_n}{SE(\hat{\theta}_n^*)} \xrightarrow{d^*} N(0,1),$$

which no longer holds when the regularity conditions do not hold. We use the bootstrap as a diagnostic test for

$$H_0 : \text{regularity conditions for Gaussian asymptotic inference hold}$$

versus

$$H_1 : \text{regularity conditions are violated.}$$

Consider a measure $\hat{d}_n$ of the distance between the bootstrap distribution $\hat{G}_n$ and Φ, the cdf of $N(0,1)$. For instance, $\hat{d}_n$ could a Kolmogorov–Smirnov (KS) type distance:

$$\hat{d}_n = \sup_{x \in \mathbb{R}} |\hat{G}_n(x) - \Phi(x)|$$

or Jarque-Bera type difference in moments:

$$\hat{d}_n = E^*(T_n^{*3}, T_n^{*4} - 3)'.$$

If bootstrap consistency holds, then (by definition) as $n \to \infty$:

$$\hat{d}_n = \sup_{x \in \mathbb{R}} |\hat{G}_n(x) - \Phi(x)| \xrightarrow{p} 0.$$

In contrast, under bootstrap failure, as in the previous three examples, it can be shown that

$$\hat{G}_n \xrightarrow{d} \mathcal{G},$$

where $\mathcal{G}(\neq \Phi)$ is a random cdf.

How can we exploit these two different limit behaviors and construct a proper test? Consider an estimator of $\hat{d}_n$ based on m bootstrap replications

$$\hat{d}_{n,m}^* = ||\hat{G}_{n,m}^* - \Phi||, \qquad \hat{G}_{n,m}^*(x) = m^{-1} \sum_{i=1}^{m} \mathbb{I}\{T_{n:i}^* \leq x\}.$$

Under the null hypothesis H_0 and depending on the distance measure,

$$\mathcal{T}_{n,m}^* = \sqrt{m} d_{n,m}^* \xrightarrow{d^*} N(0,1)$$

for $m \to \infty$ as $n \to \infty$, but such that $m/n \to 0$ (of course it converges to the absolute value of a standard normal distribution if the distance is based on KS type distance). Under the alternative, the test diverges. The hypothesis H_0 is rejected if $\mathcal{T}_{n,m}^*$ is significantly different from zero.

3 Questions

1. Consider example 1. Use histograms and 1,000 Monte Carlo (MC) simulation repetitions to investigate the non-normality of the estimator in formula (1) in the location model for $\theta \in \{0.0, 0.1, 0.2, \ldots, 0.6\}$ and $n = 100$. Explain how the distribution of $\hat{\theta}_n$ changes as θ increases.

2. Generate a table with rejection frequencies (at a significance level of $\alpha = 5\%$) for the bootstrap test in the location model, considering $\theta \in \{0.0, 0.1, 0.2, \ldots, 0.6\}$ and $n \in \{100, 400\}$. For each of the 1,000 samples, simulate a bootstrap sample and test whether it is significantly different from the normal distribution using a nominal significance level of 5% and $m = n$. For the bootstrap test, select two normality tests from the following list: Anderson-Darling, Cramér-von Mises, Shapiro-Wilk, Lilliefors, (Robust) Jarque-Bera, or Doornik–Hansen. Which test is more powerful?

3. Now consider Example 2. Investigate the non-normality of the sample mean $\hat{\theta}_n$ using Q-Q plots for $\theta_0 = 0$, $n \in \{100, 200, 400\}$, and a t-distribution with $v \in \{1, 2, 5\}$ degrees of freedom for ε_i. Identify which part of the distribution shows the largest deviation from normality.

4. Generate a table with rejection frequencies of the bootstrap test for $\theta = 0$, $n \in \{100, 200, 400\}$, and a t-distribution with $v \in \{1, 2, 5\}$ degrees of freedom for ε_i. Only consider the most powerful normality test identified in Question 2. Now, examine the following variations of m (noting that $m \to \infty$ as $n \to \infty$, but $m/n \to 0$): $n^{1/2}$, $n^{4/5}$, $n/2$, and n.

5. Finally, consider Example 3. Using 1,000 MC simulation repetitions, investigate the non-normality of the estimator in formula (2) in the weak IV model under the following conditions: the instruments are irrelevant ($\pi = 0$), weak ($\pi = 50/\sqrt{n}$), and strong ($\pi = 1000/\sqrt{n}$) for $\mathbb{C}(\varepsilon_i, u_i) = 0.7$ and $n \in \{100, 200, 400\}$. Describe the non-normality according to your preferred method.

6. Generate Q-Q plots of the conditional bootstrap distribution for $\pi = c/\sqrt{n}$, where $c \in \{0, 50, 100, 1000\}$ and $n = m = 100$. Specifically, generate four samples for each value of c, apply the (infeasible residual) bootstrap to each sample, and create a Q-Q plot. This process should result in four graphs. Describe the general patterns you observe.

7. Generate a table with rejection frequencies in the weak IV model under the following conditions: the instruments are irrelevant ($\pi = 0$), weak ($\pi = 50/\sqrt{n}$), and strong ($\pi = 1000/\sqrt{n}$) for $\mathbb{C}(\varepsilon_i, u_i) = 0.7$ and $n \in \{100, 200, 400\}$. As before, use a significance level of 5%, 1,000 MC simulations and $m = n^{4/5}$. Use your preferred test for normality.

8. Redo the simulation from Question 7, but now consider the distance from the Gaussian distribution computed over the set $[1.96, \infty)$, defined as

$$\hat{d}_n = \sup_{x \in [1.96, \infty)} |\hat{G}_n(x) - \Phi(x)|.$$

For this, you must first determine a critical value for this distance for the appropriate sample sizes in a separate simulation (again use $\alpha = 5\%$). Report the critical values and compare the rejection frequencies of the normality tests between Questions 7 and 8.

9. Until now, all simulations have assumed homoskedasticity. Suppose heteroskedasticity is present in example 3, such that $\mathbb{C}(x_i, \varepsilon_i^2) \neq 0$. Describe at least two methods to address heteroskedasticity in the bootstrap. For each method, provide a detailed explanation of the original and bootstrapped test statistic, the resampling scheme, and the rejection rule for normality. You do not need to carry out a simulation.