Monte Carlo and Bootstrap Analysis

Econ 487

Outline

- Monte Carlo analysis of estimators
- Parametric Bootstrap Analysis
- Semiparametric Bootstrap Analysis

What is Monte Carlo?

- Simulation using pseudo-random-numbers to evaluate the performance of estimators given specific data generating processes

Monte Carlo

- Specify the true DGP: e.g., R=mu+e, e~iidN
- S1. Simulate sample of T observations from DGP
- S2. Calculate estimate (i.e., mu_hat) and store
- S3. Repeat S1 and S2 for M simulations
- S4. Report summary statistics for M draws of estimator

Monte Carlo Experiment

- How does OLS do for the CER model with mu=3, sigma=5, and T=50, 200?
- Estimator for mu?
- Confidence interval for mu?

Bootstrap?

- Use sample data to inform DGP used in Monte Carlo Analysis of estimator
- E.g., S&P500 stock returns and bootstrap percentile confidence interval

Confidence Intervals

- Rule to find set of possible parameter values that includes true parameter in (1-alpha)% of samples (alpha is usually 5%)
- Standard 95% CI: mu_hat +/- t(alpha/2)*SE corresponds to inverted t-test rule (i.e., collect null hypotheses for mu that cannot be rejected using a t-test)

Percentile Approach

- Suppose we knew the distribution of mu_hat(rv) - mu
- Collect mu0's for which mu_hat(estimate) - mu0 lies within the 2.5% and 97.5% percentile of mu_hat(rv) - mu

Equivalence

- In a simple setting (e.g., truth is CER model with Normal errors), and assuming known variance, standard CI and percentile CI are the same

Problem

- In reality, we don’t know the distribution of mu_hat(rv) - mu?

Possible Solution

- Use estimated parameters for DGP in order to conduct Monte Carlo simulation to find distribution of mu_hat(rv) - mu, when mu=mu_hat(estimate)
- This is a (parametric) bootstrap experiment

Pivotal

- To the extent that the distribution of mu_hat(rv) - mu does not depend on the value of mu, the distribution is said to <u>pivotal</u>
- If it is pivotal, the bootstrap experiment will work perfectly and is, perhaps, more intuitive than inverted t

Bootstrap Experiment

- Simulate distribution of mu_hat(rv) - mu
- Sort draws from distribution
- Find 2.5th percentile and 97.5th percentile
- Use this to determine values of mu0 for which mu_hat(estimate) would not fall below 2.5th percentile or above 97.5th percentile => 95% confidence interval

Why Bootstrap?

- In this simple case, as long as the conditions for OLS are satisfied, inverted t CI works just as well as (and is easier to compute than) bootstrap percentile CI
- However, when t-distribution is not good approximation to finite sample distribution (e.g., in presence of serial correlation), bootstrap can outperform inverted t

Outperform?

- For confidence intervals, we care about actual *coverage* versus nominal coverage (i.e., 95%)
- Undercoverage is common problem for CIs
- We also want to minimize the *expected length* of the interval given correct coverage

Semiparametric Bootstrap

- In the previous experiment, we assumed that returns were Normally distributed when simulating from bootstrap DGP
- What if they are not Normal?

Draw from Empirical Distribution

- For simulations, we can draw residuals from empirical distribution (with replacement)
- Proceed as before with parametric bootstrap experiment
- For sizable T, performs as well as when returns are actually Normal (e.g., consider fake data), but provides more accurate CIs when returns are not Normal