ECON 546

Monte Carlo and Bootstrap Simulation

1, Introduction

Monte Carlo and Bootstrap simulations are rather similar, but they also differ in some important respects. In each case, the primary intention is to simulate the true sampling distribution for some statistic that is of interest to us, for problems where we cannot determine the form of this sampling distribution exactly, by analytic means. This statistic may be an estimator, or it may be a test statistic.

In the former case, knowledge of the sampling distribution, via Monte Carlo simulation, will enable us to investigate the bias, MSE *etc*. of the estimator. In the latter case knowledge of the sampling distribution, *via* Monte Carlo simulation, when the null hypothesis is true will enable us to determine the extent to which there is any "size-distortion" *(i.e.*, the extent to which the true probability of a Type I error differs from what is being assumed, on the basis of the asymptotic distribution). When the null is false, the simulated sampling distribution, obtained via Monte Carlo methods, will enable to us to explore the power of the test in question, in finite samples.

Bootstrap simulation also provides insights into the sampling distribution, and in the case of estimators it again gives us a basis for determining the bias, MSE, *etc*. In the case of tests. Bootstrap simulation provides a simple way of determining the exact p-value for the particular data that we are using.

It is important to realize that there are many variants of Monte Carlo and Bootstrap simulation. Here, we will just compare the simplest versions of each, and we will use a regression model as the vehicle for illustrating what is involved.

2. Monte Carlo Simulation

(a) Estimator Properties

Suppose that want to learn something about the bias and MSE of a particular estimator of the slope parameter in the following simple regression model:

$$y_i = \beta x_i + \varepsilon_i \quad ; \quad \varepsilon_i \sim i.i.d.\ N[0, \sigma^2] \quad ; \quad i = 1, 2, \ldots, n \tag{1}$$

The steps involved for simulating the sampling distribution of this estimator, and approximating its bias and MSE, using Monte Carlo methods, are as follows:

(i) *Assume* a value for σ^2 and 'n' and generate a random sample of 'n' values for the ε_i's from the $N[0, \sigma^2]$ distribution.

(ii) *Assume* a value for β and for the 'n' sample values of the regressor, x.

(iii) Using equation (1), generate an artificial sample of 'n' (random) values for y.

(iv) Using the actual x and artificial y sample values, estimate equation (1) using the estimator of interest, and save the value of the point estimate, $\hat{\beta}^{(1)}$.

(v) Repeat steps (iii) and (iv) many, many times, each time saving the $\hat{\beta}^{(j)}$ values, where the superscript '(j)' refers to the j^{th} repetition of the experiment.

(vi) Look at the distribution of the $\hat{\beta}^{(j)}$ values you have created (j = 1, 2, ……, N), where 'N' is the number of repetitions. **This distribution of values will approach the true sampling distribution of $\hat{\beta}$, as $N \to \infty$.**

(vii) The true bias of $\hat{\beta}$ is $E[\hat{\beta}] - \beta$, which can be approximated here by the quantity $\frac{1}{N}\sum_{j=1}^{N}\hat{\beta}^{(j)} - \beta$, which we might call the "empirical bias".

Similarly the MSE of $\hat{\beta}$ is $E[(\hat{\beta} - \beta)^2]$, which can be approximated here by the quantity $\sum_{j=1}^{N}[(\hat{\beta}^{(j)} - \beta)^2]$, which we might call the "empirical MSE".

Note: The results you obtain will be specific to the choice of n, σ^2, β and the 'n' sample values of the regressor, x, at steps (i) and (ii) above. Typically, a Monte Carlo experiment will involve an exploration of different choices of these characteristics of the problem, as the sampling distribution of $\hat{\beta}$ will usually depend on one or more of these choices.

(b) Test Properties

Similar steps are involved if we want to explore the properties of some test that we perform after estimating equation (1). Specifically, to illustrate matters, suppose that we wish to test the hypothesis H_0: $\beta = \beta_0$ *vs*. H_A: $\beta > \beta_0$. Let our test statistic be $S_n = S_n(\mathrm{y})$, and suppose that we know the asymptotic distribution of this statistic if H_0 is true. Then, our test may involve the decision: "Reject H_0 if $S_n > c(\alpha)$", where the critical value, c, is chosen to ensure a significance level of α, if the asymptotic distribution of S_n holds. That is,

$$\alpha = \int_{c}^{\infty} f_0^a(S_n(y) \mid \beta)dy$$

where $f_0^a(.)$ is the density for the *asymptotic* distribution of S_n when H_0 is true. The problem is that the *finite* sample null distribution of the test statistic may be different from its asymptotic distribution, and we may not know the form of the finite sample distribution. So, if we use the critical value, $c(\alpha)$, based on the asymptotic distribution, it probably will not correspond to a significance level of α in finite samples. That is:

$$\int_{c}^{\infty} f_0^n(S_n(y) \mid \beta)dy \neq \alpha \qquad (2)$$

where $f_0^n(.)$ is the density for the distribution of S_n when the sample size is 'n', and when H_0 is true. The extent to which the inequality holds in equation (2) is a measure of the "size distortion" associated with wrongly using the asymptotic critical value.

The steps for a Monte Carlo experiment that would determine the true significance level are as follows:

(i) Assume a value for σ^2 and 'n' and generate a random sample of 'n' values for the ε_i's from the $N[0,\sigma^2]$ distribution.

(ii) Assign the value of β_0 for β, and assign values for the 'n' sample values of the regressor, x.

(iii) Using equation (1), generate an artificial sample of 'n' (random) values for y.

(iv) Using the actual x and artificial y sample values, estimate equation (1) using the estimator of interest, and save the value of the test statistic, $S_n^{(1)}$.

(v) Repeat steps (iii) and (iv) many, many times, each saving the $S_n^{(j)}$ values, where the superscript '(j)' refers to the j^{th} repetition of the experiment.

(vi) Look at the distribution of the $S_n^{(j)}$ values you have created (j = 1, 2, ……, N), where 'N' is the number of repetitions. **This distribution of values will approach the true sampling distribution of S_n, as $N \to \infty$.**

(vii) The true significance level ("size") of the test is the proportion of times that $S_n^{(j)}$ values exceed $c(\alpha)$.

(viii) We can also impute the new finite-sample size-α critical value by looking at the distribution of the $S_n^{(j)}$ values, and choosing a "cut-off point" such that a proportion, α, of the values exceed this cut-off point. The cut-off point, $c^n(\alpha)$, will be the critical value that *should* be used in the finite-sample case.

To determine the 'raw' power of the test, based on the (wrong) asymptotic critical value, we would repeat the experiment with different values of $\beta \neq \beta_0$ at step (ii). We would end at step (vii), where the rejection rate would now be the 'power' of the test. Varying the values of $\beta \neq \beta_0$ would allow us to build up a picture of the power curve for the test.

To investigate the (size-adjusted) power of the test, which is really of more interest, we would perform the above Monte Carlo experiment to determine the power, but we would replace $c(\alpha)$ by the true finite-sample critical value, $c^n(\alpha)$.

3. Bootatrap Simulation

The main disadvantage of a Monte Carlo experiment is that the results will be specific to the choices that we make for the parameter values, the data values and the sample size. If the estimator's properties, or the properties of the test, vary as these quantities change, a large experiment will be needed in order to obtain results with any degree of generality. If we are interested in the performance of an estimator or a test, just for some particular situation that we are in, then Bootstrap simulation offers a powerful alternative way of proceeding.

(a) Estimator Properties

Again, suppose that want to learn something about the bias and MSE of a particular estimator of the slope parameter in the following simple regression model:

$$y_i = \beta x_i + \varepsilon_i \quad ; \quad \varepsilon_i \sim i.i.d.\ N[0,\sigma^2] \quad ; \quad i = 1, 2, \ldots, n \tag{3}$$

The steps involved for simulating the sampling distribution of this estimator, and approximating its bias and MSE, using the Bootstrap method, are as follows:

(i) Estimate equation (3), using our actual (n) sample values for y and x. Let the point estimator of β be $\hat{\beta}$, and let the residuals be $e_i = y_i - \hat{\beta}x_i$; i = 1, 2, ….., n.

(ii) Draw a (Bootstrap) sample on n values from the (n) e_i values, by sampling ***with replacement***. Denote the values in this first Bootstrap sample by $e_i^{(1)}$; i = 1, 2, ….,n.

(iii) Generate a sample of 'n' artificial y_i values:

$$y_i^{(1)} = \hat{\beta}x_i + e_i^{(1)} .$$

(iv) Using the actual x and artificial $y^{(1)}$ sample values, estimate equation (3) using the estimator of interest, and save the value of the point estimate, $\hat{\beta}^{(1)}$.

(v) Repeat steps (ii) to (iv) many, many times, each time saving the $\hat{\beta}^{(j)}$ values, where the superscript '(j)' refers to the j^{th} repetition of the experiment.

(vi) Look at the distribution of the $\hat{\beta}^{(j)}$ values you have created (j = 1, 2, ……, N), where 'N' is the number of repetitions. **This distribution of values will approach the true sampling distribution of $\hat{\beta}$, as $N \to \infty$.**

(vii) The true bias of $\hat{\beta}$ is $E[\hat{\beta}] - \beta$, which can be approximated here by the quantity $\sum_{j=1}^{N} \hat{\beta}^{(j)} - \hat{\beta}$, which we might call the "empirical bias".

Similarly the MSE of $\hat{\beta}$ is $E[(\hat{\beta} - \beta)^2]$, which can be approximated here by the quantity $\sum_{j=1}^{N} [(\hat{\beta}^{(j)} - \hat{\beta})^2]$, which we might call the "empirical MSE".

(b) Test Properties

Once again, similar steps are involved if we want to explore the properties of some test that we perform after estimating equation (1). Let's continue to suppose that we wish to test the hypothesis H_0: $\beta = \beta_0$ *vs*. H_A: $\beta > \beta_0$. Let our test statistic be $S_n = S_n$(y), and suppose that we know the asymptotic distribution of this statistic if H_0 is true. Then, our test may involve the decision: "Reject H_0 if $S_n > c(\alpha)$", where the critical value, c, is chosen to ensure a significance level of α, if the asymptotic distribution of S_n holds. That is,

$$\alpha = \int_c^{\infty} f_0^a(S_n(y) \mid \beta) dy$$

where $f_0^a(.)$ is the density for the *asymptotic* distribution of S_n when H_0 is true.

The steps for a Bootstrap experiment that would determine the true significance level for the test ***with our data and sample size*** are as follows:

(i) Estimate equation (3), ***imposing the conditions for the null hypothesis to be true***, using our actual (n) sample values for y and x. Let the test statistic be S_n, and let the residuals be $e_i = y_i - \beta_0 x_i$; i = 1, 2, ….., n; because the under H_0, $\beta = \beta_0$ in this very simple example. (More generally, there would be other coefficients that would be estimated.)

(ii) Draw a (Bootstrap) sample on n values from the (n) e_i values, by sampling ***with replacement***. Denote the values in this first Bootstrap sample by $e_i^{(1)}$; $i = 1, 2, \ldots, n$.

(iii) Generate a sample of 'n' artificial y_i values:

$$y_i^{(1)} = \hat{\beta}_R x_i + e_i^{(1)} .$$

where in this very simple example, the restricted estimator that imposes H_0 is just $\hat{\beta}_R = \beta_0$.

(iv) Using the actual x and artificial $y^{(1)}$ sample values, estimate equation (3) and save the value of the test statistic, $S_n^{(1)}$.

(v) Repeat steps (ii) to (iv) many, many times, each time saving the $S_n^{(j)}$ values, where the superscript '(j)' refers to the j^{th} repetition of the experiment.

(vi) Look at the distribution of the $S_n^{(j)}$ values you have created (j = 1, 2, ……, N), where 'N' is the number of repetitions. **This distribution of values will approach the true sampling distribution of S_n, as $N \rightarrow \infty$.**

(vii) The finite-sample Bootstrapped p-value is then the proportion of the N times that the $S_n^{(j)}$ values exceed the original S_n value computed at step (i).

Once again, keep in mind that there are many variations on the Bootstrap, and this is especially true when it comes to working with confidence intervals, tests, and p-values. Don't be surprised if you encounter other ways of doing this.