Assignment 1 - Bootstrap

Advanced Econometrics 2

Stepan Svoboda, Jan Hynek

January 11, 2018

1 Report

1.1 Question 1

In this question we were asked to simply simulate the coverage frequency of the first-order approximation using standard error obtained by the delta method. Our standard error obtained from the delta method is

$$se(\hat{\theta}) = \sqrt{\cos^2(\bar{X}) * var(X)/n}.$$

Using this standard error we have managed to get a coverage rate of approximately 71%.

1.2 Question 2

To determine the N for which the first-order asymptotic approximation gets the coverage frequency of the 95% CI correctly at around 95% we iterate over the Monte Carlo with the parameter N. Once we get the desired value, we stop. In our case it was $N = 1750$.

1.3 Question 3

We can answer this question by adapting the code used in the first question. If we calculate the jackknife/bootstrap bias-corrected estimate of $\hat{\theta}^{BC}_{jack/boot}$ and replace the standard $\hat{\theta}$ in the interval calculation we obtain the desired value.

We got coverage frequency 67% for the bootstrap bias-corrected estimate and 67.3% for the jackknife one. This gives us result comparable with the outcome from the first question which is expected. The difference from the first-order asymptotic approximation is that our estimate should be less biased/unbiased and the entire confidence interval just shifts somewhat. Plus we have a skewed distribution, which also should diminish the effectiveness of this approach. Based on this behavior a difference of 3 percentage points in coverage frequency in a MC study of 1000 repetitions does not pose a problem.

1.4 Question 4

Unlike in question 3 we have replaced the standard error in question 1 with a bootstrapped value and not the $\hat{\theta}$. By taking the jackknife/bootstrap value of the standard error we have managed to increase the coverage frequency to 76.7% in case of jackknife and 69.6% in case of bootstrap.

This seems to be somewhat better than both the first-order asymptotic approximation and using the bias-corrected estimate of $\hat{\theta}$ though the evidence is only scarce.

1.5 Question 5

In this question we investigate the performance of the percentile and percentile-t method. The percentile method created the confidence interval by ordering the 499 bootstrapped estimates, i.e. $\hat{\theta}^*$, and then taking the lower and upper $\alpha/2$ quantiles. This led to the coverage rate of 86.9%.

The second method is the percentile-t method which is more complex. The outer for loop in this case is the same as in the first question but the quantiles of the t-statistic are different. We calculate them using the bootstrap - we simulate the 499 $\hat{\theta}^*$ and using the delta method we also obtain for each one their appropriate standard error. Then we calculate for each of these pairs a t-statistic centered around the $\hat{\theta}$. This leads to an asymptotic refinement but in our case of small sample it gave us only a coverage frequency of 84.1%

We prefer percentile method in this case - the results are (counter-intuitively) better, and this method is much easier to implement and grasp as well. This also answers that these results are not in line with theory, as percentile-t method should have provided us with asymptotic refinement.

1.6 Question 6

This time around we take the script from the last question and augment the center of the confidence interval. We calculate the bias-corrected $\hat{\theta}_{jack/boot}$ and replace the standard $\hat{\theta}$ in the calculations. We also recalculated the original percentile-t for this randomly generated sample so these three options are fully comparable. Our results are:

- percentile-t - 83.5% CF
- percentile-t jackknife - 81.4% CF
- percentile-t bootstrap - 81.2% CF

We can see that the confidence intervals with bias-corrected centers are somewhat worse but the difference is not large for a MC study.

1.7 Question 7

In the last question that requires coding we took the code from the fifth question and entered the bias correction into the t-ratio. For every bootstrap iteration we calculated the bias using jackknife (using the code from question 4) and it gave us a very decent result of 93.9% coverage frequency. We also calculated the standard percentile-t method as in question 5 and got 83.2% coverage frequency. We can safely conclude this last option gave us by far the best results and got us into the confidence interval of 5% around the 95% CI which was shown in question 2. Thus this approach would be preferred in practice even though it is computationally by far the most demanding.

1.8 Question 8

We can use the standard paired bootstrap in this case but that does not provide any asymptotic refinement and the residual bootstrap fails in this setting. There are two ways to get the asymptotic refinement, we can either use parametric or non-parametric approach. The parametric approach entails specifying the heteroskedasticity which can be quite difficult and not always feasible thus we would prefer to use the non-parametric approach, i.e. the wild bootstrap.

The wild bootstrap leads to asymptotic refinement and does not force us to specify the heteroskedasticity. The idea behind this bootstrap is randomly transforming all of our residuals for the specific observations (we do not randomly assign the residuals to observations). These transformations can be various and the most used ones are multiplying the residual by 1 or - 1, both with probability $\frac{1}{2}$ or taking $\xi \sim \mathcal{N}(0,1)$ and multiplying the residual by ξ. There is also a more complicated version taking $-0.681\hat{u}_i$ with probability of approximately 0.7236 and $1.681\hat{u}_i$ with probability approximately 0.2764.