Advanced Econometrics 2 (2017/2018)
Bootstrap Methods - UvA
Programming Assignment
Instructions: hand in before Friday 17:00 hours, 12 January 2018 in teams of 2.

- E-mail your programming code to: N.P.A.vanGiersbergen@UvA.nl.
- You may also e-mail your solutions (max. 4 pages text; including tables) assuming basic econometric knowledge.
- MATLAB code is preferred, but other languages are allowed (however, do not use standard R-bootstrap libraries!). Clarify your code with comments and include a statement of originality (see MATLAB code).
- The purpose of this assignment is for you to gain practical experience with resampling methods.
- Do not copy work of others. This will be considered as fraud!

In this assignments, you will investigate several methods to construct confidence intervals. Suppose

$$Z_i \sim N(0, s^2) \qquad \text{and} \qquad X_i = \exp(Z_i),$$

so X_i is log-normally distributed with $\mu = E[X_i] = \exp(1/2s^2)$. We are interested in an equal-tailed two-sided confidence interval $[l, u]$ for

$$\theta = \sin(\mu)$$

such that

$$P[l \leq \theta \leq u] = 95\%.$$

1. On Blackboard, you can find MATLAB code for generating the data for $n = 50$ based on $REP = 1000$ replications. Derive the standard error based on the delta method (see for instance Section 7.2.8 of Cameron and Trivedi). Simulate the coverage frequency (cov. freq.) of the first-order asymptotic approximation, i.e. the frequency that the confidence interval (CI)

$$[\hat{\theta} - 1.96SE(\hat{\theta}), \hat{\theta} + 1.96SE(\hat{\theta})]$$

contains the true parameter θ for $s = 1.3$.

2. Determine the minimal n (in multiples of 50) such that the cov. freq. of the CI of question 1 is contained in the interval

$$0.95 \pm 1.96\sqrt{0.95 \times 0.05/1000} = [0.9365, 0.9635].$$

For the questions below, consider $n = 50$ again.

3. Next, we consider the finite-sample properties of $\hat{\theta}$. Determine the bias of $\hat{\theta}$ using simulation. Code the bias-corrections based on the jackknife and the bootstrap (499 bootstrap replications):

$$\hat{\theta}_{jack}^{BC} \qquad \text{and} \qquad \hat{\theta}_{boot}^{BC}.$$

Using simulation, investigate which method seems to be able to correct the bias most accurately. Does it also improve the cov. freq. when using a bias-corrected CI, i.e.

$$[\hat{\theta}_{jack/boot}^{BC} - 1.96SE(\hat{\theta}), \hat{\theta}_{jack/boot}^{BC} + 1.96SE(\hat{\theta})]?$$

4. Next, investigate the effect of replacing the SE based on the delta method by the jackknife and bootstrap standard errors:

$$SE_{jack}(\hat{\theta}) \qquad \text{and} \qquad SE_{boot}(\hat{\theta}).$$

So, simulate the cov. freq. of

$$[\hat{\theta} - 1.96SE_{jack/boot}(\hat{\theta}), \hat{\theta} + 1.96SE_{jack/boot}(\hat{\theta})].$$

5. Next, implement and investigate the performance of the following 95% equal-tailed two-sided bootstrap intervals: the percentile and percentile-t. Which method do you prefer? Are the results in line with the theory (you don't have to prove any claims).

6. Consider the cov. freq. of the two bootstrap intervals when you take the bias of the estimator into account, i.e. using $\hat{\theta}_{jack/boot}^{BC}$ as the center of the intervals.

7. Finally, consider bias correction in the t-ratio, i.e. take the quantiles of

$$T^* = \frac{\hat{\theta}^* - Bias_{jack}(\hat{\theta}) - \hat{\theta}}{SE(\hat{\theta}^*)}$$

for the percentile-t method (using $\hat{\theta}$ as the center). Comment on the performance and logic of this bootstrap implementation.

8. Describe (in some detail) how you would bootstrap in case of heteroskedasticity (so adding a non-constant variance to the DGP). You don't have to code (or simulate) the procedure.

Good luck!