Econometrics - Problem Set 5

Gualtiero Azzalini

Exercise 2

I first report the whole code. At the beginning I set the following values: $n = 10$ the sample size, $k = 1000$ the number of times to repeat part (b), $reps = 3000$ the number of bootstrap repetitions, $\alpha = 0.05$ for the CI, $\theta_0 = \exp(\mathbb{E}[x]) = 1$.

```
%% Exercise 2

% Set parameters
n = 10; k = 1000; reps = 3000; alpha=0.05; theta0 = 1; tic

for j=1:k
        x = randn(n,1).*sqrt(6);
        theta = exp(mean(x));
        bs=randi([1 10],n,reps);
        x_star(:,:)=x(bs(:,:));
                for i=1:reps
                        Tn(i,1) = (sqrt(n)*(exp(mean(x_star(:,i)))-theta)*...
                                        (std(x_star(:,i))*theta)^(-1))';
                end
        Tn_abs = abs(Tn);
        z(j,1) = quantile(Tn_abs,1-alpha);
        CI_bs(j,1) = theta-z(j,1)*(std(x)*theta)*sqrt(n)^(-1);
        CI_asy(j,1) = theta -1.96*(std(x)*theta)*sqrt(n)^(-1);
        CI_bs(j,2) = theta+z(j,1)*(std(x)*theta)*sqrt(n)^(-1);
        CI_asy(j,2) = theta +1.96*(std(x)*theta)*sqrt(n)^(-1);
        if rem(j/200,1) == 0
                j
        end
        if theta0<CI_bs(j,1) | theta0>CI_bs(j,2)
                dummy_bs(j,1)=0;
        else
                dummy_bs(j,1)=1;
        end
        if theta0<CI_asy(j,1) | theta0>CI_asy(j,2)
                dummy_asy(j,1)=0;
        else
```

```
                dummy_asy(j,1)=1;
        end
end
toc

%Percentage theta0 is in the interval
perc_bs = 100*sum(dummy_bs)/k;
perc_asy = 100*sum(dummy_asy)/k;
```

Part (a)

The following code does what required in the question:

```
x = randn(n,1).*sqrt(6);
theta = exp(mean(x));
```

The first line generates a random sample of $n = 10$ from a $\mathcal{N}(0,6)$ and computes $\hat{\theta} = \exp(\bar{x}_n)$ where $\bar{x}_n$ is the sample mean.

Part (b)

The following code performs what required in the question:

```
bs=randi([1 10],n,reps);
x_star(:,:)=x(bs(:,:));
                for i=1:reps
                        Tn(i,1) = (sqrt(n)*(exp(mean(x_star(:,i)))-theta)*...
                                        (std(x_star(:,i))*theta)^(-1))';
                end
Tn_abs = abs(Tn);
z(j,1) = quantile(Tn_abs,1-alpha);
CI_bs(j,1) = theta-z(j,1)*(std(x)*theta)*sqrt(n)^(-1);
CI_asy(j,1) = theta -1.96*(std(x)*theta)*sqrt(n)^(-1);
CI_bs(j,2) = theta+z(j,1)*(std(x)*theta)*sqrt(n)^(-1);
CI_asy(j,2) = theta +1.96*(std(x)*theta)*sqrt(n)^(-1);
```

In the first two lines I draw with replacement $reps = 3000$ new bootstrap samples, that i collect in the matrix x_{star}. Basically in the first line I create a 10x3000 matrix containing numbers from 1 to 10. The integers in each column represent which values in the original sample should be put in each bootstrap sample. Then I compute the t-statistic:

$$T_n(F_n) = \frac{\sqrt{n}\left(\hat{\theta}^* - \hat{\theta}\right)}{\hat{\sigma}^*}$$

Where $\hat{\theta}^* = \exp(\bar{x}_n^*)$ and $\hat{\theta} = \exp(\bar{x}_n)$ and $\hat{\sigma}^* = \sigma(x^*)\exp(\bar{x}_n) = \sigma(x^*)\hat{\theta}$. $\hat{\sigma}^*$ is computed through the delta-method and $\sigma(x^*)$ is the standard deviation of the sample. In fact, according to the delta-method, if $g(\theta)$ is a continuously differentiable function, then $\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} \mathcal{N}(0, \sigma^2 g'(\theta)^2)$ in the univariate case. In this case $g(\theta) = \exp(\theta)$ hence we can use the formula above as consistent estimator of the standard error.

Then, I take the absolute value of T_n and I compute the $1-\alpha$ quantile of this distribution to get $z_{n,\alpha}^*$ (in my case $\alpha = 0.05$). Finally I compute the CI in the bootstrap case as $\hat{\theta} \pm z_{n,\alpha}^* \frac{\sigma(x)}{\sqrt{n}}$ and in the asmptotic case as $\hat{\theta} \pm 1.96\frac{\sigma(x)}{\sqrt{n}}$. Note that I use $\sigma(x)$, the standard error of the original sample x as consistent estimator (in this case I could have used also the asymptotic variance as I know the true distribution but (i) usually the original distribution is not known (ii) the standard error of the sample is a consistent estimator and hence results are basically the same).

Part (c)

I repeat part (b) 1000 times as required and I compute the percentages as follows:

```
        if theta0<CI_bs(j,1) | theta0>CI_bs(j,2)
                dummy_bs(j,1)=0;
        else
                dummy_bs(j,1)=1;
        end
        if theta0<CI_asy(j,1) | theta0>CI_asy(j,2)
                dummy_asy(j,1)=0;
        else
                dummy_asy(j,1)=1;
        end
end
toc

%Percentage theta0 is in the interval
perc_bs = 100*sum(dummy_bs)/k;
perc_asy = 100*sum(dummy_asy)/k;
```

In the first part I define two dummies (one for the asymptotic and one for the bootstrap that is one if the true parameter $\theta_0 = 1$ is contained in the j^{th} iteration and it's zero otherwise. Then I compute the percentage of times that θ_0 is contained in the CI for the two cases. Results are in the table below:

% - Bootstrap CI	% - Asymptotic CI
94.9	87.2

Clearly the percentage is higher in the bootstrap case because, given that the t-stat is pivotal, the coverage of the bootstrap is $O(n^{-2})$ in the symmetric case, while it's $O(n^{-1/2})$ in the asymptotic case.

Exercise 3

Part (a)[1]

To verify clean maximum we can check that the conditions of Lemma 2.1 are satisfied (as the latter holds in general).

- Θ is compact by hypothesis
- $Q(\theta_0) > Q(\theta)$ for each $\theta \in \Theta$ with $\theta \neq \theta_0$: $Q(\theta)$ is a quadratic form with W positive definite and symmetric which implies that is minimized at zero, i.e. when $\gamma(\theta_0) = g_0$. Since by assumption we know that there is a unique θ_0 such that $\gamma(\theta_0) = g_0$ then this condition holds.
- $Q : \Theta \to \mathbb{R}$ is continuous: we know by assumption that $\gamma : \Theta \to \mathbb{R}^k$ is continuous hence, to prove that Q is continuous we have to show that $(g_0 - \gamma(\theta))$ is continuous in θ. Consider any $\theta^* \in \Theta$ and let $(\theta_n)_{n\in\mathbb{N}} \subset \Theta$ be a sequence such that $||\theta_n - \theta^*|| \to 0$ as $n \to \infty$. Then since $\gamma : \Theta \to \mathbb{R}^k$ is continuous: $\lim_{n\to\infty}(g_0 - \gamma(\theta_n)) = g_0 - \gamma(\theta^*)$ and since algebraic operations are continuous, then this proves that the whole function Q is continuous.

As all the conditions required by the lemma are satisfied clean maximum holds.

Part (b)

We want to show $\sup_{\theta\in\Theta} |Q_n(\theta) - Q(\theta)| = o_p(1)$.

Now, consider the following:

$$2(Q(\theta) - Q_n(\theta)) = (g_n - \gamma(\theta))'\hat{W}(g_n - \gamma(\theta)) - (g_0 - \gamma(\theta))'W(g_0 - \gamma(\theta)).$$

Add and subtract $(g_n - \gamma(\theta))'W(g_n - \gamma(\theta))$:

$$\Rightarrow (g_n - \gamma(\theta))'(\hat{W} - W)(g_n - \gamma(\theta)) + (g_n - \gamma(\theta))'W(g_n - \gamma(\theta)) - (g_0 - \gamma(\theta))'W(g_0 - \gamma(\theta)).$$

Note that the last two terms can be written as:

$$(g_n - g_0)'W(g_n - g_0) - \gamma(\theta)'W(g_n - g_0) - (g_n - g_0)'W\gamma(\theta).$$

Therefore:

$$\begin{aligned}\sup_{\theta\in\Theta} 2\,|Q_n(\theta) - Q(\theta)| \leq & \sup_{\theta\in\Theta} \left|(g_n - \gamma(\theta))'(\hat{W} - W)(g_n - \gamma(\theta))\right| \\ & + \sup_{\theta\in\Theta} |(g_n - g_0)'W(g_n - g_0)| \\ & + \sup_{\theta\in\Theta} |\gamma(\theta)'W(g_n - g_0)| \\ & + \sup_{\theta\in\Theta} |(g_n - g_0)'W\gamma(\theta)|\,.\end{aligned}$$

Now note the following:

[1]Note: I assume first and second moments of $g(w_t)$ are finite

$$
\begin{aligned}
g_n - g_0 = \tfrac{1}{n}\textstyle\sum_{t=1}^{n} g(w_t) - \mathbb{E}[g(w_t)] = o_{a.s.}(1) & \qquad by\, Ergodic\, Theorem \\
||\hat{W} - W|| = o_p(1) & \qquad by\, assumption \\
\sup_{\theta\in\Theta} (||\gamma(\theta)||) < \infty & \qquad by\, Weierstrass\, as\, \gamma(\theta)\, continuous\, and\, \Theta\, compact \\
\sup_{\theta\in\Theta} ||g_n - \gamma(\theta)|| < \infty & \qquad by\, the\, above\, and\, finite\, moments\, assumption
\end{aligned}
$$

and that g_n and g_0 don't depend on θ, so that:

$$
\begin{aligned}
\sup_{\theta\in\Theta} 2\,|Q_n(\theta) - Q(\theta)| \leq & \left(\sup_{\theta\in\Theta} ||g_n - \gamma(\theta)||\right)^2 ||\hat{W} - W|| \\
& + (||g_n - g_0||)^2\, ||W|| \\
& + \sup_{\theta\in\Theta} (||\gamma(\theta)||)\, ||W||\, ||g_n - g_0|| \\
& + ||g_n - g_0||\, ||W|| \sup_{\theta\in\Theta} (||\gamma(\theta)||) \\
& = O_p(1) o_p(1) + o_p(1) * constant + O_p(1) * constant * o_p(1) + o_p(1) * constant * O_p(1) \\
& = o_p(1).
\end{aligned}
$$

This verifies uniform convergence.

Part (c)

By part (a) and part (b) and by the fact that we are dealing with an extremum estimator, using Theorem 2.1 we conclude that $\hat{\theta} \xrightarrow{p} \theta_0$ as $n \to \infty$.