**Economics 628** **Term 1 2017/2018**

**Topics in Applied Econometrics I** **Hiro Kasahara**

# Homework 4

(Due: Monday, October 30 at the *start* of the class)

Note: Study groups *discussing* the problems are strongly encouraged. But please write your own answers and submit your own programs (no copy and paste from your classmate's program!).

## 1 Bootstrap Confidence Intervals

In this exercise, we compute bootstrap confidence intervals for the estimates in Homework 1 and 2. We use $B = 999$ bootstrap replications.[1]

1. Compute percentile-t equal-tailed bootstrap confidence interval for $\hat{\theta}_1$ and $\hat{\theta}_2$, corresponding to the estimated coefficients of 'nearc4' and 'IQ', for the model of college decision in Homework 1 using "HW1_data" as follows.

Step 1: For each $b$, draw a bootstrap sample $\{(y_1^{*(b)}, x_1^{*(b)}), ..., (y_N^{*(b)}, x_N^{*(b)})\}$ with replacement from the original data $\{(y_1, x_1), ..., (y_N, x_N)\}$ by repeating the following procedure for $i = 1, ..., N$: (i) Draw $u_i$ ∼Uniform$[0, 1]$ using 'rand' command. **(ii) If** $u_i < 1/N$**, set** $(y_i^{*(b)}, x_i^{*(b)}) = (y_1, x_1)$**. Else if** $u_i \geq (N-1)/N$**, then set** $(y_i^{*(b)}, x_i^{*(b)}) = (y_N, x_N)$**. Otherwise, if** $(j-1)/N < u_i \leq j/N$**, then set** $(y_i^{*(b)}, x_i^{*(b)}) = (y_j, x_j)$**.** Note, in consequence, a bootstrap sample $\{(y_1^{*(b)}, x_1^{*(b)}), ..., (y_N^{*(b)}, x_N^{*(b)})\}$ will have multiple values.

Step 2: Using the $b$-th bootstrap sample, $\{(y_1^{*(b)}, x_1^{*(b)}), ..., (y_N^{*(b)}, x_N^{*(b)})\}$, estimate the coefficient $\theta$ in model (1) in Homework 1 and their standard errors. Denote the estimate and the standard errors by $\hat{\theta}_k^{*(b)}$ and $s(\hat{\theta}_k^{*(b)})$ for $k = 1, 2$. Compute bootstrap-t statistics by $T_k^{*(b)} = (\hat{\theta}_k^{*(b)} - \hat{\theta}_k)/s(\hat{\theta}_k^{*(b)})$.

[1] You probably want to start from some small number, say, $B = 20$ to make sure that your code is working. When you estimate the parameter with bootstrap samples, use the original estimate under the original sample as an initial value for optimization routine. This is because the bootstrap estimate is within $\sqrt{N}$ neighborhood of the original estimate. In fact, theoretically, taking 3 or 4 Newton-Raphson steps will give a valid bootstrap confidence interval.

Step 3: Repeat Step 1 for $b = 1, ..., B$.

Step 4: Denote the $\alpha$-th sample quantile of the simulated statistics $\{T_k^{*(1)}, ..., T_k^{*(B)}\}$ by $\hat{q}_k^*(\alpha)$. Sort $\{T_k^{*(1)}, ..., T_k^{*(B)}\}$ from the smallest to the largest using the Matlab command "sort". With **B=999**, $\hat{q}_k^*(\alpha/2)$ and $\hat{q}_k^*(1-\alpha/2)$ for $\alpha = 0.05$ can be computed by the 25-th and 975-th smallest values among $\{T_k^{*(1)}, ..., T_k^{*(B)}\}$, respectively.[2]

Step 5: Compute the percentile-t bootstrap confidence interval for $k = 1, 2$ as:

$$C_3 = [\hat{\theta}_k - s(\hat{\theta}_k)\hat{q}_k^*(1-\alpha/2), \quad \hat{\theta}_k - s(\hat{\theta}_k)\hat{q}_k^*(\alpha/2)].$$

2. Compute also two different versions of percentile bootstrap confidence intervals for $\alpha = 0.05$ as follows.

   (a) Repeat Steps 1-4 in Question 1 but using $T_k^{*(b)} = \hat{\theta}_k^{*(b)}$ in place of $T_k^{*(b)} = (\hat{\theta}_k^{*(b)} - \hat{\theta}_k)/s(\hat{\theta}_k^{*(b)})$. Then, in Step 5, compute the 95 percentile bootstrap confidence interval for $k = 1, 2$ as:

   $$C_1 = [\hat{q}_k^*(\alpha/2), \quad \hat{q}_k^*(1-\alpha/2)].$$

   (b) Repeat Steps 1-4 in Question 1 but using $T_k^{*(b)} = \hat{\theta}_k^{*(b)} - \hat{\theta}_k$ in place of $T_k^{*(b)} = (\hat{\theta}_k^{*(b)} - \hat{\theta}_k)/s(\hat{\theta}_k^{*(b)})$. Then, compute the percentile bootstrap confidence interval for $k = 1, 2$ as:

   $$C_2 = [\hat{\theta}_k - \hat{q}_k^*(1-\alpha/2), \quad \hat{\theta}_k - \hat{q}_k^*(\alpha/2)].$$

   Which of these three bootstrap CIs do you recommend *least*? Briefly discuss the advantages and the disadvantages of the bootstrap confidence intervals, $C_1$, $C_2$, and $C_3$. [For your reference, see chapter 10 of Hansen, B. (2016).]

3. Compute the bootstrap confidence intervals, $C_1$, $C_2$, and $C_3$, with $\alpha = 0.05$ for the average partial effect of IQ on college attendance probability in Homework 1.

4. It is also possible to do "parameteric bootstrap" when the model you estimate is a parametric bootstrap by drawing the bootstrap data from the estimated parametric model. The only

[2] **We choose $B$ so that $v = (B+1)(1-\alpha)$ is an integer, where the $v$-th value will give $(1-\alpha)$ quantile.**

difference between non-parametric and parametric bootstrap in this context is that, in Step 1, the bootstrap data is generated from the estimated parameter model as follows:

Step 1: For each $b$, draw a bootstrap sample $\{(y_1^{*(b)}, x_1^{*(b)}), ..., (y_N^{*(b)}, x_N^{*(b)})\}$ as follows. For each $i = 1, 2, ..., n$, draw $\epsilon_i^{*(b)}$ from $N(0, 1)$, conditional on $x_i$ (which is the value of $x_i$ in the original data), let $y_i^{*(b)} = 1(x_i'\hat{\theta} + \epsilon_i^{*(b)})$. Set $x_i^{*(b)} = x_i$ for all $b$. That is, $(y_i^{*(b)}, x_i^{*(b)}) = (1(x_i'\hat{\theta} + \epsilon_i^{*(b)}), x_i)$ for $i = 1, ..., n$.

Compute the bootstrap confidence intervals, $C_1$, $C_2$, and $C_3$, with $\alpha = 0.05$ for the average partial effect of IQ on college attendance probability in Homework 1 based on the parametric bootstrap.

## 2 Maximum Simulated Likelihood for Dynamic Discrete Choice Models

Consider a dynamic discrete choice model of Homework 2. Suppose that, conditional on observable variables, $c_i$ is iid drawn from $N(0, \sigma^2)$ instead of two point mixture distribution. This assignment asks you to estimate the parameter $(\alpha, \rho, \sigma)'$ by simulated maximum likelihood using the data set "HW2_data". Denote the $(T+1) \times N$ data matrix by $Y = [y_1, ..., y_N]$ with $y_i = (y_{i0}, ..., y_{iT})'$, where $T = 6$ and $n = 718$. Try $R = 500$ first and then try $R = 1000$ to see check if the estimates are not so sensitive to the choice of $R$ once $R$ is sufficiently large.

1. Draw $\eta_i^r$ from $N(0, 1)$ for $r = 1, ..., R$ and $i = 1, ..., n$, and store them into a $R \times n$ matrix $Q = (\eta_1, ..., \eta_n)$ with $\eta_i = (\eta_i^1, ..., \eta_i^R)'$. Note that $\eta_i^r$'s are drawn only once and they are fixed throughout the optimization routine. Try also antithetics by drawing $\eta_i^r$ from $N(0, 1)$ for $r = 1, ..., R/2$ and set $\eta_i = (\eta_i^1, ..., \eta_i^{R/2}, -\eta_i^1, ..., -\eta_i^{R/2})'$.

2. Given that $c \sim N(0, \sigma^2)$, the log likelihood function is given by

$$\mathcal{L}(\alpha, \rho, \sigma^2 | Y) = \sum_{i=1}^{N} \ln L_i(\alpha, \rho, \sigma^2 | y_i), \tag{1}$$

where

$$L_i(\alpha, \rho, \sigma^2 | y_i) = \int \left( \Pr(y_{i0}|\alpha, \rho, c) \prod_{t=1}^{T} \Pr(y_{it}|y_{i,t-1}, c) \right) (1/\sigma)\phi(c/\sigma) dc,$$

whereas $\Pr(y_{i0}|\alpha, \rho, c)$ and $\Pr(y_{it}|y_{i,t-1}, c)$ are given by equations (1) and (3) in Homework 2. The value of $L_i(\alpha, \rho, \sigma|y_i)$ can be approximately simulated as:

$$\check{L}_i(\alpha, \rho, \sigma^2 | y_i, Q) = (1/R) \sum_{r=1}^{R} \Pr(y_{i0}|\alpha, \rho, \sigma\eta_i^r) \prod_{t=1}^{T} \Pr(y_{it}|y_{i,t-1}, \sigma\eta_i^r), \tag{2}$$

where $\sigma\eta_i^r$ is the $r$-th simulated value of $c$. The log of simulated likelihood function is given by:

$$\check{\mathcal{L}}(\alpha, \rho, \sigma^2 | Y) = \sum_{i=1}^{N} \ln \left( (1/R) \sum_{r=1}^{R} \left( \Pr(y_{i0}|\alpha, \rho, \sigma\eta_i^r) \prod_{t=1}^{T} \Pr(y_{it}|y_{i,t-1}, \sigma\eta_i^r) \right) \right) \tag{3}$$

Write a function m-file that computes the negative value of the log of simulated likelihood (3) given the value of $\theta = (\alpha, \rho, \sigma)$, the data matrix $Y$, and the simulated vector $Q$ as its input. Estimate the parameter $\theta$ for the data set "HW2_data" and compute the standard errors of $(\alpha, \rho, \sigma)$.

# 3 MSM

Suppose that we are interested in estimating the mean of random variable $Y$, i.e., $\theta = E[Y]$. Suppose that the distribution of $Y$ is given by $N(\theta, 1)$. Let $\{Y_i\}_{i=1}^n$ be the data set, where each observation is independently drawn from $N(\theta, 1)$. We would like to estimate $\theta$ by the Method of Simulated Moments (MSM) given $n \times R$ simulated random draws from $N(0, 1)$, $\{\{\eta^{i(r)}\}_{r=1}^R\}_{i=1}^n$. We consider the asymptotics where $n \to \infty$ while $R$ is fixed.

1. What is the MSM estimator?

2. Derive the asymptotic distribution of the MSM estimator.