Econ 312: Problem Set 1

Professor Magne Mogstad

Due Tuesday, April 14 before the lecture

Problem 1

Consider the nonparametric Roy model:

$$Y = DY_1 + (1 - D)Y_0$$

$$D = \mathbb{1}[U \leq \nu(Z)]$$

$$(Y_0, Y_1, U) \perp\!\!\!\perp Z$$

$$\mathbb{E}[Y_d^2] < \infty, d \in \{0, 1\}$$

Where (Y_0, Y_1, U) are unobserved.

a) What assumptions about distribution of U did you see in class? Are they restrictive?

b) Define MTE and derive ATE, ATT and ATUT as weighted averages of MTE.

c) Interpret the weights you received in the previous part.

Problem 2

Consider the following set up for the Roy model:

$$
\begin{gathered}
Y_1 = U_1 \\
Y_0 = U_0 \\
\begin{pmatrix} U_1 \\ U_0 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix}\right) \\
D = \mathbb{1}[U_1 > U_0] \\
Y = DY_1 + (1-D)Y_0 = \underbrace{(Y_1 - Y_0)}_{\beta} D + Y_0
\end{gathered}
$$

a) Derive the expression for β_{OLS}. What treatment effect does it correspond to in the case $D \perp\!\!\!\perp (Y_1, Y_0)$?

b) Derive the expressions for ATT and ATUT, comment on their relative magnitudes and signs?

c) What is ATE in this case?

d) Derive $\frac{\partial ATT}{\partial \rho}$, $\frac{\partial ATUT}{\partial \rho}$, $\frac{\partial \beta_{OLS}}{\partial \rho}$, provide intuitive explanation for your results. Derive $\frac{\partial ATT}{\partial \sigma}$, $\frac{\partial ATUT}{\partial \sigma}$, $\frac{\partial \beta_{OLS}}{\partial \sigma}$, is there any simple intuitive explanation for these results?

e) Set: $\sigma = 2$ and $\rho = 0.5$. Draw $N = 10000$ pairs (U_0, U_1) and compute ATE, ATT, ATUT and β_{OLS}. (Note, you have all the counterfactuals in this set up - take advantage of this fact). Compute $\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0]$, what parameter does it correspond to? Repeat the exercise for $\sigma = 2$, $\rho = 0$ and $\sigma = 2$, $\rho = -0.5$. Fix $\rho = 0.5$ and vary σ to verify your conclusions from **d)**.

f) Claim: in this set up $D \perp\!\!\!\perp (Y_1, Y_0) \iff \rho = 0$.

Argue whether this claim is correct or not. Support your conclusion with the results from **e)**.

Problem 3

The purpose of this exercise is to introduce you to bootstrapping procedure.

In previous classes we learned that under i.i.d. assumption we can receive the following results:

$$\hat{\beta} \xrightarrow{p} \bar{\beta} = \mathbb{E}[X_i X_i']^{-1}\mathbb{E}[X_i Y_i]$$

$$\sqrt{N}(\hat{\beta} - \bar{\beta}) \xrightarrow{d} \mathcal{N}(0, V)$$

$$\exists \hat{V} \xrightarrow{p} V$$

$$\text{se}(\hat{\beta}_k) = \sqrt{\frac{1}{N}\text{diag}(\hat{V})_k}$$

Monte Carlo simulations

Consider the model:

$$Y_i = X_i'\beta + U_i$$

$$U_i | X_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$$

a) Define $\beta = (2,3)^T, \sigma^2 = 4$; generate $N = 10000$ values for $X \in \mathbb{R}^2$(constant and one more covariate). Using your value for σ^2 draw U's (you can make it completely independent of X). Finally, compute Y's. Estimate $\hat{\beta}$ and its standard errors from your data using standard OLS formulas.

b) Using X, β and σ^2 from part **a)** draw $S = 10000$ of $\underbrace{U^{(s)}}_{N\times 1 \text{ vector for each s}}$ and corresponding $Y^{(s)}$. For each $Y^{(s)}$ compute:

$$\hat{\beta}^{(s)} = (\sum_{i=1}^{N} X_i X_i')^{-1}(\sum_{i=1}^{N} X_i Y_i^{(s)})$$

$$\sqrt{\widehat{\text{Var}}[\hat{\beta}_k^{(s)} | X_1, \ldots, X_N]} = \sqrt{\frac{1}{S}\sum_{s=1}^{S}(\hat{\beta}_k^{(s)})^2 - \left(\frac{1}{S}\sum_{s=1}^{S}\hat{\beta}_k^{(s)}\right)^2} \underbrace{\xrightarrow{p}}_{?} \text{se}(\hat{\beta}_k | X_1, \ldots, X_N)$$

Justify the "?" step in the last line. Plot a histogram for the first component of $\beta^{(s)}$.

Nonparametric Bootstrap

Roughly, the idea behind the bootstrap procedure is that we expect large sample of observed data to behave similarly to the population. Then, following the logic of Monte Carlo simulations, we want to draw samples from this population and conduct the inference (so we do not have to specify the data generating process as we did in the previous part).

Let us work through an example. Consider RCT set up:

$$Y = DY_1 + (1 - D)Y_0$$

$$D \perp\!\!\!\perp (Y_1, Y_0)$$

a) Define constant values for $Y_1 = 5 + U_1$ and $Y_0 = 2 + U_2$, where U_1 and U_2 are independent standard normal variables, assign $D \in \{0, 1\}$ randomly (define probability with which $D = 1$, 0.5 would be a good choice) to $N = 10000$ individuals.

Note, we can rewrite the initial outcome equation:

$$Y = Y_0 + D \underbrace{(Y_1 - Y_0)}_{\beta}$$

Estimate β using a standard OLS procedure, compute the standard errors, argue that OLS gives consistent coefficient estimates.

b) Now, we will create bootstrap samples. From the initial data that you generated draw $N = 10000$ pairs (Y_i, D_i) choosing each of the original data pairs with probability $\frac{1}{N}$ (without replacement). Repeat this procedure a total of $S = 10000$ times. Now you should have $S = 10000$ samples of $N = 10000$ observations each generated from the original sample. Repeat the computations from the **Monte Carlo** part and compute $\sqrt{\widehat{\text{Var}}[\hat{\beta}^{(s)}]}$, plot histogram for $\hat{\beta}^{(s)}$. What would happen if you drew Y and D independently from the original sample, instead of as a pair?(hint: think about the value of coefficient on D)