Ec2147, Spring 2021

Professor Bryan Graham

Problem Set 1: part (a)

Due: March 16th, 2021

Problem sets are due by 5PM EST. You may work in groups (indeed I encourage you to all work together), but each student should turn in their own write-up.

Helpful reading: Newey (1990), Brown and Newey (1998) and Graham (2011). For an alternative, and highly elegant approach, see Chamberlain (1987).

Let $Z = (X^{\prime}; Y)^{\prime}$ be a vector of modeling variables and $\{Z_i\}_{i=1}^{N}$ an independent and identically distribution random sequence drawn from unknown distribution F_0. The sole prior restriction on F_0 is that, for some known function $\quad(Z; \ )$ and unknown parameter $\ _0 \in \mathrm{B} \quad \mathrm{R}^K$

$$\mathrm{E}[\quad(Z; \ _0)] = 0.$$

Assume that $\dim(\quad(Z; \ _0)) = J > \dim(\ _0) = K.$

Let $a(Z; \ )$ be some known function of Z and $\ $. The goal is to e$\square$ciently estimate the mean

$$\ _0 = \mathrm{E}[a(Z; \ _0)].$$

1. Let $X = \ X^{\prime}; W\ '$ with W a binary treatment indicator (also partition $\quad = \ \tilde{}; \ ^{\prime} \ ^{\prime}$) and assume that

$$\Pr(Y = 1 | X) = \quad X^{\prime}\tilde{} + W \quad .$$

Let $\ _l = \mathrm{E}[Y | X \in \mathcal{X}_l]$ for $l = 1, \ldots, L$. Assume that these conditional means are known to the econometrician (e.g., from register data). The target estimand is

$$\ _0 = \mathrm{E}\left[\quad X^{\prime}\tilde{} + \quad\quad X^{\prime}\tilde{} \ \right].$$

Outline a plausible empirical setting to which you could adapt the above components. Show how your setting is accommodated by the general setup outlined above.

2. Calculate the semiparametric variance bound, say $I(\ _0)^{-1}$, for $\ _0$. You may use either the approach outlined in Newey (1990) (and also lecture) or that of Chamberlain (1987).

3. Show that the semiparametric variance bound, say $I(\ _0)^{-1}$, for $\ _0$ is

$$I(\ _0)^{-1} = \quad_{aa} \quad a \quad _0^{-1} \ ^{\prime}_a + \quad a \quad _0^{-1} \ _0 \quad \ _0 \ I(\ _0)^{-1} \ ^{\prime}_0 \ _0^{-1} \ ^{\prime}_a \quad \ ^{\prime}_0$$

where $\ _{aa} = \mathrm{E}[(a(Z; \ _0) \quad \ _0)(a(Z; \ ) \quad \ _0)^{\prime}]$; $\ _a = \mathrm{E}[(a(Z; \ _0) \quad \ _0) \quad (Z; \ _0)^{\prime}]$, $\ _0 = \mathrm{E}[\quad(Z; \ _0) \quad (Z; \ _0)^{\prime}]$, $\ _0 = \mathrm{E}\left[\frac{\partial \quad(Z; \ _0)}{\partial \ ^{\prime}}\right]$ and $\ _0 = \mathrm{E}\left[\frac{\partial a(Z; \ _0)}{\partial \ ^{\prime}}\right]$. As before, you may use either the approach of Chamberlain (1987) or Newey (1990) (in this case I also recommend Brown and Newey (1998)). Interpret each of the three components of the inverse information expression above.

4. Assume that $\hat{\ }$ is an e$\square$cient estimate of $\ _0$.

(a) Is $\hat{\ } = \frac{1}{N}\sum_{i=1}^{N} a\ Z_i; \hat{\ }$ e$\square$cient? Explain.