203B Question Bank

John Friedman

April 7, 2024

1. (Comp Fall 2023 B1) Let $\{Y_i, X_i\}_{i=1}^n$ be an iid sample with $Y_i \in R$ and $X_i \in R^d$ satisfying:

$$Y_i = X_i'\beta_0 + \epsilon_i \text{ and } E[(\epsilon - E[\epsilon])(X - E[X])] = 0 \quad (1)$$

In what follows please be careful that we have assumed that the covariance between ϵ and $X = 0$, but not necessarily that $E[\epsilon X] = 0$.

(a) Does model (1) imply that $E[\epsilon X] = 0$? Prove your answer.

Solution: No. Taking the expectation of both sides of the moment restriction, we get:

$$\begin{aligned} E[(\epsilon - E[\epsilon])(X - E[X])] &= 0 \\ E[\epsilon X - \epsilon E[X] - E[\epsilon]X + E[\epsilon]E[X]] &= 0 \\ E[\epsilon X] - E[\epsilon E[X]] - E[E[\epsilon]X] + E[E[\epsilon]E[X]] &= 0 \\ E[\epsilon X] - E[\epsilon]E[X] - E[\epsilon]E[X] + E[\epsilon]E[X] &= 0 \\ E[\epsilon X] - E[\epsilon]E[X] &= 0 \end{aligned}$$

Thus, we see that $E[\epsilon X] = E[\epsilon]E[X]$, which is not necessarily equal to zero.

(b) Show that under our assumptions we must have the restriction:

$$E[\{(Y_i - E[Y]) - (X_i - E[X])'\beta_0\}(X - E[X])] = 0$$

Solution:

(c) Can the moment restriction in part (b) identify β_0? if X_i contains a constant. Why or why not? Justify your answer.

(d) Let $\bar{Y}_n \equiv \frac{1}{n}\sum_{i=1}^n (Y_i)$ and $\bar{X}_n \equiv \frac{1}{n}\sum_{i=1}^n (X_i)$, and define the estimator:

$$\hat{\beta}_n = \arg\min_{b \in R^d} \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y}_n - (X_i - \bar{X}_n)'b)^2$$

Establish the consistency of $\hat{\beta}_n$ to β_0 clearly stating any assumptions you make.

(e) Another researcher is concerned that you did not include a constant in (1). He instead prefers the traditional regression:

$$(\tilde{\alpha}_n, \tilde{\beta}_n) \equiv \arg\min_{a \in R, b \in R^d} \frac{1}{n}\sum_{i=1}^n (Y_i - a - X_i'b)^2$$

How does his estimator $\tilde{\beta}_n$ compare to your estimator $\hat{\beta}_n$ from part (c)? Justify your answer. Hint: Frisch-Waugh-Lovell Theorem.

2. (Comp Fall 2023 B2) Let $Z \in 0, 1$ be an instrument, $D \in 0, 1$ be a treatment, and Y an observable outcome. Throughout, assume a LATE framework in which there are two potential outcome $(Y(0), Y(1))$, two potential treatment assignments $(D(0), D(1))$, and assume that the observable D and Y are determined according to:

$$D = D(0) + Z(D(1) - D(0))$$

i.e. we observe the potential outcome corresponding to the actual treatment status, and the potential treatment assigment corresponding to the realization of Z. Further, assume that $Y(0), Y(1), D(0), D(1)$ are all independent of Z. The monotonocity condition that:

$$P(D(1) \geq D(0)) = 1$$

and that we have available an iid sample $\{Y_i, D_i, Z_i\}_{i=1}^n$ of (Y, D, Z).

(a) Show that under the stated assumptions we must have

$$P(D = 1|Z = 1) = P(D(1) = 1) \text{ and } P(D = 1|Z = 0) = P(D(0) = 1)$$

(b) use part (a) to argue that if the monotonicity assumption (e.g. $P(D(1) \geq D(0)) = 1)$), is correct, then the following restriction must hold:

$$P(D = 1|Z = 1) - P(D = 1|Z = 0) \geq 0$$

(c) In order to check whether the monotonocity assumption is correct, we compute the following sample analague to the quantities in $P(D = 1|Z = 1) - P(D = 1|Z = 0) \geq 0$:

$$\frac{\sum_{i=1}^n (D_i Z_i)}{\sum_{i=1}^n (Z_i)} - \frac{\sum_{i=1}^n (D_i(1 - Z_i))}{\sum_{i=1}^n (1 - Z_i)}$$

(recall that $D \in 0, 1$ and $Z \in 0, 1$). Carefully derive the asymptotic distribution of the estimator above.

(d) Propose an estimator for the asymptotic variance of the estimator in part (c). You do not need to formally establish consistency.

(e) Use the results of parts (b)-(d) to propose a test of the monotonocity assumption $P(D(1) \geq D(0)) = 1$. You do not need to formally establish results, but you should clearly outline exactly how to computer the test if we want the probability of a Type I error to be α.