1 Conditional Expectations

Suppose X and Z are binary random variables, while Y is just a generic random variable.

Problem 1.1. Is $E[Y|X] = BLP[Y|X]$?

Solution. If we show that CEF is linear in the conditioning variable, then we can conclude that CEF = BLP. The provided statement is always true since

$$\mathbb{E}[Y|X] = X\mathbb{E}[Y|X=1] + (1-X)\mathbb{E}[Y|X=0] = \mathbb{E}[Y|X=0] + X[\mathbb{E}[Y|X=1] - \mathbb{E}[Y|X=0]]$$

■

Problem 1.2. Is $E[Y|X,Z] = BLP[Y|X,Z]$?

Solution. Note that we can write:

$$\begin{aligned}\mathbb{E}[Y|X,Z] = \mathbb{E}[Y|0,0] &+ X(\mathbb{E}[Y|1,0] - \mathbb{E}[Y|0,0]) \\ &+ Z(\mathbb{E}[Y|0,1] - \mathbb{E}[Y|0,0]) + XZ(\mathbb{E}[Y|1,1] - \mathbb{E}[Y|1,0] - \mathbb{E}[Y|0,1] + \mathbb{E}[Y|0,0])\end{aligned}$$

which means that we need the XZ term to get CEF. So No. ■

Problem 1.3. Is $E[Y|X,Z,XZ] = BLP[Y|X,Z,XZ]$?

Solution. Yes, by the argument provided above. ■

Problem 1.4. If $E[Y|Z] = 0$ and $E[Y|X] = 0$, then is $E[Y|X,Z] = 0$?

Solution. We will provide a counterexample. Let X and Z be independent Bernoulli random variables with mean 0.5, and let

$$Y = \mathbb{I}(X=Z) - 0.5$$

Then we have

$$\mathbb{E}[Y|X=1] = P(Z=1)(1-0.5) + P(Z=0)(0.-0.5) = 0 = \mathbb{E}[Y|X=0]$$

and thus

$$\mathbb{E}[Y|Z] = 0, \mathbb{E}[X|Z] = 0$$

but

$$\mathbb{E}[Y|X,Z] \neq 0$$

■

2 True or False

Consider the following two regressions:

$$
\begin{aligned}
[1]: Y &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + U \\
[2]: Y &= \beta_0^* + \beta_1^* X_1 + \beta_2^* X_2 + U^*
\end{aligned}
$$

Is the following statement true or false? "If you have a regressor X_1 that is only correlated to X_3 through X_2, then β_1 should be the same."

Solution. It is true.

▷ In regression [1], consider the partitioned regression:

$$Y - BLP\left(Y|1, X_2\right) = \beta_1 \left(X_1 - BLP\left(X_1|1, X_2\right)\right) + \beta_3 \left(X_3 - BLP\left(X_3|1, X_2\right)\right) + \left(U - BLP\left(U|1, X_2\right)\right)$$

▷ In regression [2], consider the partitioned regression

$$Y - BLP\left(Y|1, X_2\right) = \beta_1^* \left(X_1 - BLP\left(X_1|1, X_2\right)\right) + \left(U^* - BLP\left(U^*|1, X_2\right)\right)$$

▷ From regression [2], we then have

$$
\begin{aligned}
\beta_1^* &= \frac{\operatorname{Cov}\left[Y - BLP\left(Y|1, X_2\right), \left(X_1 - BLP\left(X_1|1, X_2\right)\right)\right]}{\operatorname{Var}\left[X_1 - BLP\left(X_1|1, X_2\right)\right]} \\
&= \beta_1 + \frac{\operatorname{Cov}\left[\beta_3 \left(X_3 - BLP\left(X_3|1, X_2\right)\right) + \left(U - BLP\left(U|1, X_2\right)\right), \left(X_1 - BLP\left(X_1|1, X_2\right)\right)\right]}{\operatorname{Var}\left[X_1 - BLP\left(X_1|1, X_2\right)\right]}
\end{aligned}
$$

▷ Since X_1 is only correlated to X_3 through X_2, it follows that

$$\operatorname{Cov}\left(X_3, \left(X_1 - BLP\left(X_1|1, X_2\right)\right)\right) = 0$$

▷ From the property of BLP, we also have:

$$
\begin{aligned}
\operatorname{Cov}\left(BLP\left(X_3|1, X_2\right), X_1 - BLP\left(X_1|1, X_2\right)\right) &= 0 \\
\operatorname{Cov}\left(BLP\left(U|1, X_2\right), X_1 - BLP\left(X_1|1, X_2\right)\right) &= 0
\end{aligned}
$$

▷ From the assumption about U, we have:

$$\operatorname{Cov}\left(U, X_1 - BLP\left(X_1|1, X_2\right)\right) = 0$$

Therefore, $\beta_1^* = \beta_1$. ■

3 Estimator for Median

We have Y_i and D_i where $Y_i = Y_i(1)D_i + Y_i(0)(1 - D_i)$. Furthermore, we had the independence assumption that $Y_i(0)$ and $Y_i(1)$ are independent of D_i. We are asked to come up with the estimator $F_1(t)$ and $F_0(t)$ where

$$F_d(t) = Pr\{Y_i(d) \leq t\}$$

The estimator for F_d must be a function of Y_i and D_i but not of $Y_i(0)$ or $Y_i(1)$.

Problem 3.1. Propose an estimator for F_d.

Solution. We only observe Y_i and D_i but not $Y_i(1)$ or $Y_i(0)$. Therefore, for $F_1(t) = Pr\{Y_i(1) \leq t\}$, we will have to use Y_is that have $D_i = 1$. Therefore, we can count among those $D_i = 1$ that have $Y_i \leq t$. Mathematically, this amounts to:

$$\hat{F}_{dn}(t) = \frac{\sum I(Y_i \leq t) I(D_i = d)}{\sum I\{D_i = d\}}$$

■

Problem 3.2. Show that the estimator is consistent.

Solution. Note that from the proposed expression:

$$\frac{\sum I(Y_i \leq t) I(D_i = d)}{\sum I\{D_i = d\}} = \frac{\frac{1}{N}\sum I(Y_i \leq t) I(D_i = d)}{\frac{1}{N}\sum I\{D_i = d\}}$$

we have

$$\frac{1}{N}\sum I(Y_i \leq t) I(D_i = d) \xrightarrow{p} P(Y_i \leq t, D_i = d)$$
$$\frac{1}{N}\sum I\{D_i = d\} \xrightarrow{p} P(D_i = d)$$

By the continuous mapping theorem, we have

$$\frac{\sum I(Y_i \leq t) I(D_i = d)}{\sum I\{D_i = d\}} \xrightarrow{p} P(Y_i \leq t | D = d) = P(Y_i(d) \leq t)$$

■

Problem 3.3. Propose an estimator for the median.

Solution. Since we have random assignment, we can use $D_i = d$ subsample to estimate all the population quantities of $Y_i(d)$. Therefore, the estimate of median for $D_i = d$ is

$$\hat{\theta}_{dn} = \inf\left(t : \hat{F}_{dn}(t) \geq 0.5\right)$$

■

Problem 3.4. What assumption do you need to make?

Solution. We need to assume that the population median set of $D = d$ to be unique, i.e.

$$F_d(\theta_d + \epsilon) > 0.5, \forall \epsilon > 0$$

■

Problem 3.5. Prove that it is consistent.

Solution. We will show that for a given $\epsilon > 0$, we have

$$P\left(\left|\hat{\theta}_{dn} - \theta_d\right| > \epsilon\right) \to 0$$

as $n \to \infty$.

▷ Case 1: $\theta_{dn} > \theta_d + \epsilon$ (which means $F_{dn}(\theta_d + \epsilon) < 0.5)$. Then:

$$\begin{aligned} P(\theta_{dn} > \theta_d + \epsilon) &\leq P(F_{dn}(\theta_d + \epsilon) < 0.5) \\ &= P(F_{dn}(\theta_d + \epsilon) - F_d(\theta_d + \epsilon) < 0.5 - F_d(\theta_d + \epsilon)) \\ &= P(F_{dn}(\theta_d + \epsilon) - F_d(\theta_d + \epsilon) < -\delta_1) \end{aligned}$$

for some $\delta_1 > 0$ by the uniqueness assumption of θ_d. Since $F_{dn}(t) \xrightarrow{p} F_d(t)$ for any t, we have that the above expression converges to zero as $n \to \infty$.

▷ Case 2: $\theta_{dn} < \theta_d - \epsilon$ (which means $F_{dn}(\theta_d - \epsilon) \geq 0.5)$. We can apply the same steps as above.

Therefore, we conclude that $\theta_{dn} \xrightarrow{p} \theta_d$. ■

4 Instruments

Consider the model $Y = X\beta + U$ and we had an instrument Z such that $E[ZU] = 0$ and $E[ZX^T]$ has full rank.

Problem 4.1. State the conditions for the instrument variables.

Solution. We require:

- ▷ Relevance: $\mathbb{E}[ZX']$ to have rank $k+1 \leq l+1$
- ▷ Exogeneity: $\mathbb{E}[ZU] = 0$
- ▷ Monotonicity: $P\left(D_1^i \geq D_0^i\right) = 1$

■

Problem 4.2. Provide an estimator for β.

Solution. The standard two-stage least squares estimator works here:

$$\hat{\beta}_n = \left(\sum \Pi_n' Z_i Z_i' \Pi_n\right)^{-1} \left(\sum \Pi_n' Z_i Y_i\right)$$

where

$$\Pi_n = \left(\sum Z_i Z_i'\right)^{-1} \left(\sum Z_i X_i'\right)$$

■

Problem 4.3. You are given two sets of datas: $(Z_i^A, Y_i^A)_{i=1}^N$ and, independently of the first sample, $(Z_i^B, X_i^B)_{i=1}^N$. How does the estimator look like?

Solution. We can consider the following estimator:

$$\tilde{\beta}_n = \left(\sum Z_i^B \left(X_i^B\right)'\right)^{-1} \sum \left(Z_i^A Y_i^A\right)$$

Consistency follows by standard argument. ■

Problem 4.4. Does the estimator above have the same asymptotic distribution as in (2)?

Solution. Note that this estimator does not have the same asymptotic distribution. To see this, write:

$$\tilde{\beta}_n = \left(\sum Z_i^B \left(X_i^B\right)'\right)^{-1} \sum \left(Z_i^A X_i^A\right) \beta + \left(\sum Z_i^B \left(X_i^B\right)'\right)^{-1} \sum \left(Z_i^A U_i^A\right)$$

which implies:

$$\sqrt{n}\left(\tilde{\beta}_n - \beta\right) = \underbrace{\sqrt{n}\left(\sum Z_i^B \left(X_i^B\right)'\right)^{-1} \sum \left(Z_i^A X_i^A - Z_i^B X_i^B\right) \beta}_{[1]} + \sqrt{n}\left(\sum Z_i^B \left(X_i^B\right)'\right)^{-1} \sum \left(Z_i^A U_i^A\right)$$

in the standard derivation, [1] is zero but here it's not. Therefore, we cannot conclude that the estimator has the same asymptotic distribution as before. ■