Solutions to Exercises of Chapter 6

October 27, 2025
Ting-Chih Hung
nuit0815@gmail.com

Problem 6.7 (Lin's estimator for covariate adjustment)

Consider two ordinary least squares (OLS) problems:

$$\min_{\gamma_1,\beta_1} \sum_{i=1}^{n} Z_i\big(Y_i - \gamma_1 - \beta_1^\top X_i\big)^2 \quad \text{and} \quad \min_{\gamma_0,\beta_0} \sum_{i=1}^{n} (1 - Z_i)\big(Y_i - \gamma_0 - \beta_0^\top X_i\big)^2.$$

The solutions are $\hat{\gamma}_1, \hat{\beta}_1$ and $\hat{\gamma}_0, \hat{\beta}_0$. Define the estimator

$$\begin{aligned} \hat{\tau}(\hat{\beta}_1, \hat{\beta}_0) &= \frac{1}{n_1} \sum_{i=1}^{n} Z_i\Big(Y_i - \hat{\beta}_1^\top X_i\Big) - \frac{1}{n_0} \sum_{i=1}^{n} (1 - Z_i)\Big(Y_i - \hat{\beta}_0^\top X_i\Big) \\ &= \hat{\gamma}_1 - \hat{\gamma}_0. \end{aligned} \tag{6.4}$$

We are asked to show the Proposition 6.2:

> The estimator $\hat{\tau}(\hat{\beta}_1, \hat{\beta}_0)$ in Equation (6.4) equals the coefficient of Z_i in the OLS fit of Y_i on $(1, Z_i, X_i, Z_i \times X_i)$, which is Lin's estimator, $\hat{\tau}_{\mathrm{L}}$.

Let's write the OLS problem of regressing Y_i on $(1, Z_i, X_i, Z_i \times X_i)$:

$$\min_{\alpha_0,\alpha_1,\alpha_2,\alpha_3} \sum_{i=1}^{n} \big(Y_i - \alpha_0 - \alpha_1 Z_i - \alpha_2^\top X_i - \alpha_3^\top (Z_i X_i)\big)^2.$$

It is equivalent to the following problem:

$$\min_{\alpha_0,\alpha_1,\alpha_2,\alpha_3} \sum_{i=1}^{n} Z_i\big(Y_i - (\alpha_0 + \alpha_1) - (\alpha_2 + \alpha_3)^\top X_i\big)^2 + \sum_{i=1}^{n} (1 - Z_i)\big(Y_i - \alpha_0 - \alpha_2^\top X_i\big)^2.$$

Let $\gamma_1 = \alpha_0 + \alpha_1$, $\beta_1 = \alpha_2 + \alpha_3$, $\gamma_0 = \alpha_0$, and $\beta_0 = \alpha_2$. Then the above problem becomes

$$\min_{\gamma_1,\beta_1,\gamma_0,\beta_0} \sum_{i=1}^{n} Z_i\big(Y_i - \gamma_1 - \beta_1^\top X_i\big)^2 + \sum_{i=1}^{n} (1 - Z_i)\big(Y_i - \gamma_0 - \beta_0^\top X_i\big)^2.$$

This is exactly a reparametrization of the two separate OLS problems we started with. Therefore, the solutions are related by

$$\begin{aligned} \hat{\gamma}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1, \\ \hat{\beta}_1 &= \hat{\alpha}_2 + \hat{\alpha}_3, \\ \hat{\gamma}_0 &= \hat{\alpha}_0, \\ \hat{\beta}_0 &= \hat{\alpha}_2, \end{aligned}$$

and we have

$$\hat{\tau}(\hat{\beta}_1, \hat{\beta}_0) = \hat{\gamma}_1 - \hat{\gamma}_0 = (\hat{\alpha}_0 + \hat{\alpha}_1) - \hat{\alpha}_0 = \hat{\alpha}_1 = \hat{\tau}_{\mathrm{L}}.$$

Problem 6.8 (Predictive and projective estimators)

We can predict $Y_i(z)$ based on X_i using the data from the treatment group ($Z_i = 1$) and control group ($Z_i = 0$) separately,

$$\hat{\mu}_1(X_i) = \hat{\gamma}_1 + \hat{\beta}_1^\top X_i,$$
$$\hat{\mu}_0(X_i) = \hat{\gamma}_0 + \hat{\beta}_0^\top X_i,$$

where $(\hat{\gamma}_1, \hat{\beta}_1)$ and $(\hat{\gamma}_0, \hat{\beta}_0)$ are the solutions to the two OLS problems in Problem 6.7.

If we use the predictions to impute the missing potential outcomes, then we can estimate the average treatment effect (ATE) by

$$\hat{\tau}_{\text{pred}} = \frac{1}{n}\left\{\sum_{Z_i=1} Y_i + \sum_{Z_i=0} \hat{\mu}_1(X_i) - \sum_{Z_i=1} \hat{\mu}_0(X_i) - \sum_{Z_i=0} Y_i\right\}, \tag{6.7}$$

which is called the *predictive estimator*.

Instead of predicting the missing potential outcomes, if we predict all potential outcomes even if they are observed, then we have the *projective estimator*:

$$\hat{\tau}_{\text{proj}} = \frac{1}{n}\sum_{i=1}^{n} \{\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\}. \tag{6.9}$$

We are asked to show that $\hat{\tau}_{\text{pred}}$ equals Lin's estimator $\hat{\tau}_{\text{L}}$. Furthermore, $\hat{\tau}_{\text{proj}}$ also agrees with $\hat{\tau}_{\text{L}}$.

We first show $\hat{\tau}_{\text{proj}} = \hat{\tau}_{\text{L}}$. By substituting the definitions of $\hat{\mu}_1(X_i)$ and $\hat{\mu}_0(X_i)$,

$$\begin{aligned}
\hat{\tau}_{\text{proj}} &= \frac{1}{n}\sum_{i=1}^{n} \left\{(\hat{\gamma}_1 - \hat{\gamma}_0) + (\hat{\beta}_1 - \hat{\beta}_0)^\top X_i\right\} \\
&= (\hat{\gamma}_1 - \hat{\gamma}_0) + (\hat{\beta}_1 - \hat{\beta}_0)^\top \bar{X} \\
&= \hat{\tau}_{\text{L}} + (\hat{\beta}_1 - \hat{\beta}_0)^\top \bar{X} \\
&= \hat{\tau}_{\text{L}},
\end{aligned}$$

where the last equality holds because of the centering of X_i (i.e., $\bar{X} = 0$).

By normal equations of the two OLS problems, we have

$$\sum_{Z_i=z} Y_i = \sum_{Z_i=z} \left(\hat{\gamma}_z + \hat{\beta}_z^\top X_i\right) = \sum_{Z_i=z} \hat{\mu}_z(X_i).$$

Plug in the above equation into Equation (6.7), we have

$$\begin{aligned}
\hat{\tau}_{\text{pred}} &= \frac{1}{n}\left\{\sum_{Z_i=1} \hat{\mu}_1(X_i) + \sum_{Z_i=0} \hat{\mu}_1(X_i) - \sum_{Z_i=1} \hat{\mu}_0(X_i) - \sum_{Z_i=0} \hat{\mu}_0(X_i)\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n} \{\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\} \\
&= \hat{\tau}_{\text{proj}} \\
&= \hat{\tau}_{\text{L}}.
\end{aligned}$$

Problem 6.11 (Regression adjustment / post-stratification of CRE)

Suppose that we have discrete covariates $C_i \in \{1, 2, \dots, K\}$. We create $K-1$ centered dummy variables

$$X_i = \left\{\mathbf{1}\{C_i = 1\} - \pi_{[1]}, \dots, \mathbf{1}\{C_i = K-1\} - \pi_{[K-1]}\right\},$$

where $\pi_{[k]} = P(C_i = k)$ for $k = 1, \dots, K$. We are asked to show that Lin's estimator $\hat{\tau}_{\mathrm{L}}$ based on X_i equals the post-stratified estimator $\hat{\tau}_{\mathrm{PS}}$:

$$\hat{\tau}_{\mathrm{PS}} = \sum_{k=1}^{K} \hat{\pi}_{[k]} \hat{\tau}_{[k]}.$$

Define $\beta_z = (\beta_{z,1}, \dots, \beta_{z,K-1})^\top$ for $z \in \{0, 1\}$. Let the predicted potential outcomes be

$$\hat{\mu}_1(X_i) = \hat{\gamma}_1 + \hat{\beta}_1^\top X_i,$$
$$\hat{\mu}_0(X_i) = \hat{\gamma}_0 + \hat{\beta}_0^\top X_i,$$

where $(\hat{\gamma}_1, \hat{\beta}_1)$ and $(\hat{\gamma}_0, \hat{\beta}_0)$ are the regression coefficients from the regression of Y_i on $(1, X_i)$ in the treatment group and control group, respectively. Note that both regressions are saturated models within each stratum, so the predicted potential outcomes equal the stratum means:

$$\hat{\mu}_1(X_i) = \hat{\bar{Y}}_{[k]}(1) \quad \text{if } C_i = k,$$
$$\hat{\mu}_0(X_i) = \hat{\bar{Y}}_{[k]}(0) \quad \text{if } C_i = k.$$

Thus,

$$\begin{aligned}
\hat{\tau}_{\mathrm{L}} &= \frac{1}{n} \sum_{i=1}^{n} \{\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\} \\
&= \frac{1}{n} \sum_{k=1}^{K} \sum_{C_i = k} \left\{\hat{\bar{Y}}_{[k]}(1) - \hat{\bar{Y}}_{[k]}(0)\right\} \\
&= \sum_{k=1}^{K} \frac{n_{[k]}}{n} \hat{\tau}_{[k]} \\
&= \sum_{k=1}^{K} \hat{\pi}_{[k]} \hat{\tau}_{[k]} \\
&= \hat{\tau}_{\mathrm{PS}}.
\end{aligned}$$

Problem 6.12 (More on the difference-in-difference estimator in the CRE)

Acronyms

ATE average treatment effect. 2
OLS ordinary least squares. 1, 2