

Solutions to Exercises of Chapter 4

September 15, 2025
Ting-Chih Hung
nuit0815@gmail.com

Problem 4.1 (Proof of Lemma 4.1)

Let $\bar{Y}(z) = \frac{1}{n} \sum_{i=1}^n Y_i(z)$ be the finite population mean under treatment z . Let $\tau_i = Y_i(1) - Y_i(0)$ be the individual treatment effect and $\bar{\tau} = \frac{1}{n} \sum_{i=1}^n \tau_i$ be the finite population average treatment effect.

We are asked to show that

$$2S(1, 0) = S^2(1) + S^2(0) - S^2(\tau),$$

where

$$\begin{aligned} S^2(z) &= \frac{1}{n-1} \sum_{i=1}^n (Y_i(z) - \bar{Y}(z))^2, \quad z \in \{0, 1\}, \\ S^2(\tau) &= \frac{1}{n-1} \sum_{i=1}^n (\tau_i - \bar{\tau})^2, \\ S(1, 0) &= \frac{1}{n-1} \sum_{i=1}^n (Y_i(1) - \bar{Y}(1))(Y_i(0) - \bar{Y}(0)). \end{aligned}$$

To see this, we first note that

$$\tau_i - \bar{\tau} = (Y_i(1) - \bar{Y}(1)) - (Y_i(0) - \bar{Y}(0)).$$

Thus, we have

$$\begin{aligned} S^2(\tau) &= \frac{1}{n-1} \sum_{i=1}^n (\tau_i - \bar{\tau})^2 \\ &= \frac{1}{n-1} \sum_{i=1}^n ((Y_i(1) - \bar{Y}(1)) - (Y_i(0) - \bar{Y}(0)))^2 \\ &= \frac{1}{n-1} \sum_{i=1}^n [(Y_i(1) - \bar{Y}(1))^2 + (Y_i(0) - \bar{Y}(0))^2 - 2(Y_i(1) - \bar{Y}(1))(Y_i(0) - \bar{Y}(0))] \\ &= S^2(1) + S^2(0) - 2S(1, 0), \end{aligned}$$

which completes the proof.

Alternative proof of Theorem 4.1

Problem 4.3 (Neymanian inference and OLS)

Consider the ordinary least squares (OLS) estimator defined as

$$(\hat{\alpha}, \hat{\beta}) = \operatorname{argmin}_{a, b} \sum_{i=1}^n (Y_i - a - bZ_i)^2.$$

We are asked to show that

$$\hat{\beta} = \hat{\tau}. \quad (4.3)$$

To see this, we first note that

$$\sum_{i=1}^n (Y_i - a - bZ_i)^2 = \sum_{Z_i=1} (Y_i - a - b)^2 + \sum_{Z_i=0} (Y_i - a)^2.$$

Taking the derivatives with respect to a and b , we have

$$\begin{aligned} \frac{\partial}{\partial a} \sum_{i=1}^n (Y_i - a - bZ_i)^2 &= -2 \sum_{Z_i=1} (Y_i - a - b) - 2 \sum_{Z_i=0} (Y_i - a), \\ \frac{\partial}{\partial b} \sum_{i=1}^n (Y_i - a - bZ_i)^2 &= -2 \sum_{Z_i=1} (Y_i - a - b). \end{aligned}$$

Setting these derivatives to zero and solving the resulting equations, we have

$$\begin{aligned} \hat{\alpha} &= \frac{1}{n_0} \sum_{Z_i=0} Y_i = \hat{Y}(0), \\ \hat{\beta} &= \frac{1}{n_1} \sum_{Z_i=1} Y_i - \hat{\alpha} = \hat{Y}(1) - \hat{Y}(0) = \hat{\tau}, \end{aligned}$$

which shows Equation (4.3).

Next, we are asked to show that the usual variance estimator from the OLS equals

$$\begin{aligned} \hat{V}_{\text{OLS}} &\equiv \left[\hat{\sigma}^2 \left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \right]_{(2,2)} \\ &= \frac{n(n_1 - 1)}{(n - 2)(n_1 - n_0)} \hat{S}^2(1) + \frac{n(n_0 - 1)}{(n - 2)(n_0 - n_1)} \hat{S}^2(0) \\ &\approx \frac{\hat{S}^2(1)}{n_0} + \frac{\hat{S}^2(0)}{n_1}, \end{aligned} \quad (4.4)$$

where $X_i = (1, Z_i)^\top$, $\hat{\sigma}^2 = (n - 2)^{-1} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}Z_i)^2$, $[\cdot]_{(2,2)}$ denotes the second diagonal element of a matrix, and

$$\hat{S}^2(z) = \frac{1}{n_z - 1} \sum_{Z_i=z} (Y_i - \hat{Y}(z))^2, \quad z \in \{0, 1\}.$$

We first compute $\hat{\sigma}^2$:

$$\begin{aligned} \hat{\sigma}^2 &\equiv \frac{1}{n - 2} \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}Z_i)^2 && \text{(Definition)} \\ &= \frac{1}{n - 2} \left[\sum_{Z_i=1} (Y_i - \hat{Y}(1))^2 + \sum_{Z_i=0} (Y_i - \hat{Y}(0))^2 \right] && \text{(Use Equation (4.3))} \end{aligned}$$

$$= \frac{n_1 - 1}{n - 2} \hat{S}^2(1) + \frac{n_0 - 1}{n - 2} \hat{S}^2(0). \quad (\text{Definition})$$

Next, we compute $(\sum_{i=1}^n X_i X_i^\top)^{-1}$:

$$\left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} = \left(\sum_{i=1}^n \begin{pmatrix} 1 & Z_i \\ Z_i & Z_i^2 \end{pmatrix} \right)^{-1} \quad (\text{Definition of } X_i)$$

$$= \begin{pmatrix} n & n_1 \\ n_1 & n_1 \end{pmatrix}^{-1} \quad (Z_i \text{ is binary})$$

$$= \begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_1} \\ -\frac{1}{n_0} & \frac{1}{n_1} \end{pmatrix}. \quad (\text{Matrix inversion})$$

Combining these results, we arrive at Equation (4.4):

$$\hat{V}_{\text{OLS}} \equiv \left[\hat{\sigma}^2 \left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \right]_{(2,2)} \quad (\text{Definition})$$

$$= \hat{\sigma}^2 \cdot \frac{n}{n_1 n_0} \quad (\text{Extract the } (2,2) \text{ element})$$

$$= \frac{n(n_1 - 1)}{(n - 2)(n_1 - n_0)} \hat{S}^2(1) + \frac{n(n_0 - 1)}{(n - 2)(n_0 - n_1)} \hat{S}^2(0). \quad (\text{Use the expression of } \hat{\sigma}^2)$$

Then, we need to show that the Eicker-Huber-White (EHW) variance estimator equals

$$\begin{aligned} \hat{V}_{\text{EHW}} &\equiv \left[\left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \left(\sum_{i=1}^n \hat{\varepsilon}_i^2 X_i X_i^\top \right) \left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \right]_{(2,2)} \\ &= \frac{\hat{S}^2(1)}{n_1} \frac{n_1 - 1}{n_1} + \frac{\hat{S}^2(0)}{n_0} \frac{n_0 - 1}{n_0} \\ &\approx \frac{S^2(1)}{n_1} + \frac{S^2(0)}{n_0}, \end{aligned} \quad (4.5)$$

where $\hat{\varepsilon}_i = Y_i - \hat{\alpha} - \hat{\beta} Z_i$.

We first compute $\sum_{i=1}^n \hat{\varepsilon}_i^2 X_i X_i^\top$:

$$\begin{aligned} &\sum_{i=1}^n \hat{\varepsilon}_i^2 X_i X_i^\top \\ &= \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta} Z_i)^2 \begin{pmatrix} 1 & Z_i \\ Z_i & Z_i^2 \end{pmatrix} \quad (\text{Definition of } X_i \text{ and } \hat{\varepsilon}_i) \end{aligned}$$

$$= \sum_{Z_i=1}^n (Y_i - \hat{Y}(1))^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \sum_{Z_i=0}^n (Y_i - \hat{Y}(0))^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (\text{Use Equation (4.3)})$$

$$= (n_1 - 1) \hat{S}^2(1) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + (n_0 - 1) \hat{S}^2(0) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad (\text{Definition of } \hat{S}^2(z))$$

$$= \begin{pmatrix} (n_1 - 1)\hat{S}^2(1) + (n_0 - 1)\hat{S}^2(0) & (n_1 - 1)\hat{S}^2(1) \\ (n_1 - 1)\hat{S}^2(1) & (n_1 - 1)\hat{S}^2(1) \end{pmatrix}. \quad (\text{Matrix addition})$$

Then, we have

$$\begin{aligned} & \hat{V}_{\text{EHW}} \\ & \equiv \left[\left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \left(\sum_{i=1}^n \hat{\varepsilon}_i^2 X_i X_i^\top \right) \left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \right]_{(2,2)} \\ & = \left[\begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_0} \\ -\frac{1}{n_0} & \frac{n}{n_1 n_0} \end{pmatrix} \begin{pmatrix} (n_1 - 1)\hat{S}^2(1) + (n_0 - 1)\hat{S}^2(0) & (n_1 - 1)\hat{S}^2(1) \\ (n_1 - 1)\hat{S}^2(1) & (n_1 - 1)\hat{S}^2(1) \end{pmatrix} \begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_0} \\ -\frac{1}{n_0} & \frac{n}{n_1 n_0} \end{pmatrix} \right]_{(2,2)} \\ & = \left(-\frac{1}{n_0} \right) \left[(n_0 - 1)\hat{S}^2(0) + (n_1 - 1)\hat{S}^2(1) \right] \left(-\frac{1}{n_0} \right) + \left(\frac{n}{n_1 n_0} \right) (n_1 - 1)\hat{S}^2(1) \left(-\frac{1}{n_0} \right) \\ & \quad + \left(-\frac{1}{n_0} \right) (n_1 - 1)\hat{S}^2(1) \left(\frac{n}{n_1 n_0} \right) + \left(\frac{n}{n_1 n_0} \right) (n_1 - 1)\hat{S}^2(1) \left(\frac{n}{n_1 n_0} \right) \\ & = \frac{(n_0 - 1)\hat{S}^2(0)}{n_0^2} + \frac{(n_1 - 1)\hat{S}^2(1)}{n_0^2} - \frac{2n(n_1 - 1)\hat{S}^2(1)}{n_1 n_0^2} + \frac{n^2(n_1 - 1)\hat{S}^2(1)}{n_1^2 n_0^2} \\ & = \frac{(n_0 - 1)\hat{S}^2(0)}{n_0^2} + \frac{(n_1^2 - 2nn_1 + n^2)(n_1 - 1)\hat{S}^2(1)}{n_1^2 n_0^2} \\ & = \frac{(n_0 - 1)\hat{S}^2(0)}{n_0^2} + \frac{(n_1 - 1)\hat{S}^2(1)}{n_1^2}, \end{aligned}$$

which shows Equation (4.5).

Finally, we are asked to show that the HC2 variance estimator exactly recovers \hat{V} , where the HC2 variance estimator is defined as

$$\hat{V}_{\text{HC2}} \equiv \left[\left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \left(\sum_{i=1}^n \frac{\hat{\varepsilon}_i^2}{(1 - h_{ii})} X_i X_i^\top \right) \left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \right]_{(2,2)},$$

with $h_{ii} = X_i^\top (\sum_{i=1}^n X_i X_i^\top)^{-1} X_i$, and

$$\hat{V} = \frac{\hat{S}^2(1)}{n_1} + \frac{\hat{S}^2(0)}{n_0}.$$

Notice that for $Z_i = 0$,

$$h_{ii} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_0} \\ -\frac{1}{n_0} & \frac{n}{n_1 n_0} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{n_0},$$

and for $Z_i = 1$,

$$h_{ii} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_0} \\ -\frac{1}{n_0} & \frac{n}{n_1 n_0} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{n_0} - \frac{1}{n_0} + \frac{n}{n_1 n_0} = \frac{1}{n_1}.$$

We can compute $\sum_{i=1}^n (1 - h_{ii})^{-1} \hat{\varepsilon}_i^2 X_i X_i^\top$:

$$\begin{aligned}
& \sum_{i=1}^n \frac{\hat{\varepsilon}_i^2}{1 - h_{ii}} X_i X_i^\top \\
&= \sum_{Z_i=0} \frac{\hat{\varepsilon}_i^2}{1 - \frac{1}{n_0}} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \sum_{Z_i=1} \frac{\hat{\varepsilon}_i^2}{1 - \frac{1}{n_1}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (\text{Use the expression of } h_{ii}) \\
&= \frac{n_0}{n_0 - 1} \sum_{Z_i=0} \hat{\varepsilon}_i^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{n_1}{n_1 - 1} \sum_{Z_i=1} \hat{\varepsilon}_i^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (\text{Factor out constants}) \\
&= n_0 \hat{S}^2(0) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + n_1 \hat{S}^2(1) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (\text{Definition of } \hat{S}^2(z)) \\
&= \begin{pmatrix} n_0 \hat{S}^2(0) + n_1 \hat{S}^2(1) & n_1 \hat{S}^2(1) \\ n_1 \hat{S}^2(1) & n_1 \hat{S}^2(1) \end{pmatrix}. \quad (\text{Matrix addition})
\end{aligned}$$

Applying similar calculations as those for \hat{V}_{EHW} , we have

$$\begin{aligned}
& \hat{V}_{\text{HC2}} \\
&\equiv \left[\left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \left(\sum_{i=1}^n \frac{\hat{\varepsilon}_i^2}{(1 - h_{ii})} X_i X_i^\top \right) \left(\sum_{i=1}^n X_i X_i^\top \right)^{-1} \right]_{(2,2)} \\
&= \left[\begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_0} \\ -\frac{1}{n_0} & \frac{n}{n_1 n_0} \end{pmatrix} \begin{pmatrix} n_0 \hat{S}^2(0) + n_1 \hat{S}^2(1) & n_1 \hat{S}^2(1) \\ n_1 \hat{S}^2(1) & n_1 \hat{S}^2(1) \end{pmatrix} \begin{pmatrix} \frac{1}{n_0} & -\frac{1}{n_0} \\ -\frac{1}{n_0} & \frac{n}{n_1 n_0} \end{pmatrix} \right]_{(2,2)} \\
&= \frac{\hat{S}^2(0)}{n_0} + \frac{\hat{S}^2(1)}{n_1},
\end{aligned}$$

which exactly takes the form Treatment effect heterogeneity of \hat{V} .

Problem 4.4 (Treatment effect heterogeneity)

Problem 4.5 (A better bound of the variance formula)

We are asked to show that

$$\text{Var}(\hat{\tau}) \leq \frac{1}{n} \left\{ \sqrt{\frac{n_0}{n_1}} S(1) + \sqrt{\frac{n_1}{n_0}} S(0) \right\}^2. \quad (4.6)$$

We derive two intermediate results.

First, since we can express $\hat{\tau}$ as

$$\begin{aligned}
\hat{\tau} &\equiv \frac{1}{n_1} \sum_{Z_i=1} Y_i - \frac{1}{n_0} \sum_{Z_i=0} Y_i \quad (\text{Definition}) \\
&= \frac{1}{n_1} \sum_{i=1}^n Z_i Y_i(1) - \frac{1}{n_0} \sum_{i=1}^n (1 - Z_i) Y_i(0) \quad (\text{Consistency})
\end{aligned}$$

$$= \sum_{i=1}^n Z_i \left\{ \frac{Y_i(1)}{n_1} + \frac{Y_i(0)}{n_0} \right\} - \frac{1}{n_0} \sum_{i=1}^n Y_i(0). \quad (\text{Rearrange terms})$$

Therefore, the variance of $\hat{\tau}$ is given by

$$\begin{aligned} & \text{Var}(\hat{\tau}) \\ &= \text{Var} \left\{ \sum_{i=1}^n Z_i \left(\frac{Y_i(1)}{n_1} + \frac{Y_i(0)}{n_0} \right) \right\} \quad (\text{Constant doesn't matter}) \\ &= n_1^2 \text{Var} \left\{ \frac{1}{n_1} \sum_{i=1}^n Z_i \left(\frac{Y_i(1)}{n_1} + \frac{Y_i(0)}{n_0} \right) \right\} \quad (\text{Factor out } n_1^2) \\ &= \frac{n_1 n_0}{n(n-1)} \sum_{i=1}^n \left(\frac{Y_i(1)}{n_1} + \frac{Y_i(0)}{n_0} - \frac{1}{n} \sum_{j=1}^n \left(\frac{Y_j(1)}{n_1} + \frac{Y_j(0)}{n_0} \right) \right)^2 \quad (\text{Lemma C.2}) \\ &= \frac{n_1 n_0}{n(n-1)} \sum_{i=1}^n \left(\frac{Y_i(1) - \bar{Y}(1)}{n_1} + \frac{Y_i(0) - \bar{Y}(0)}{n_0} \right)^2 \quad (\text{Definition of } \bar{Y}(z)) \\ &= \frac{n_1 n_0}{n(n-1)} \left[\frac{1}{n_1^2} \sum_{i=1}^n (Y_i(1) - \bar{Y}(1))^2 + \frac{1}{n_0^2} \sum_{i=1}^n (Y_i(0) - \bar{Y}(0))^2 \right. \\ &\quad \left. + \frac{2}{n_1 n_0} \sum_{i=1}^n (Y_i(1) - \bar{Y}(1))(Y_i(0) - \bar{Y}(0)) \right] \quad (\text{Expand the square}) \\ &= \frac{n_0}{nn_1} S^2(1) + \frac{n_1}{nn_0} S^2(0) + \frac{2}{n} S(1, 0). \quad (\text{Definition of } S^2(z) \text{ and } S(1, 0)) \end{aligned}$$

Second, applying the Cauchy inequality, we have

$$\begin{aligned} & |S(1, 0)| \\ &\equiv \left| \frac{1}{n-1} \sum_{i=1}^n (Y_i(1) - \bar{Y}(1))(Y_i(0) - \bar{Y}(0)) \right| \quad (\text{Definition}) \\ &\leq \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i(1) - \bar{Y}(1))^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i(0) - \bar{Y}(0))^2} \quad (\text{Cauchy inequality}) \\ &= S(1)S(0), \quad (\text{Definition}) \end{aligned}$$

which further implies that

$$-S(1)S(0) \leq S(1, 0) \leq S(1)S(0).$$

Applying these two results, we have

$$\begin{aligned} & \text{Var}(\hat{\tau}) \\ &= \frac{n_0}{nn_1} S^2(1) + \frac{n_1}{nn_0} S^2(0) + \frac{2}{n} S(1, 0) \quad (\text{Use the above result}) \\ &\leq \frac{n_0}{nn_1} S^2(1) + \frac{n_1}{nn_0} S^2(0) + \frac{2}{n} S(1)S(0) \quad (\text{Use the above result}) \end{aligned}$$

$$= \frac{1}{n} \left\{ \sqrt{\frac{n_0}{n_1}} S(1) + \sqrt{\frac{n_1}{n_0}} S(0) \right\}^2, \quad (\text{Complete the square})$$

which shows Equation (4.6).

Problem 4.6 (Vector version of Neyman (1923))

Problem 4.7 (Inference in the BRE)

Acronyms

EHW	Eicker-Huber-White. 3
HC	heteroskedasticity-consistent. 4
OLS	ordinary least squares. 1 , 2