Solutions to Exercises of Chapter 11

October 25, 2025
Ting-Chih Hung
nuit0815@gmail.com

Problem 11.1 (Another version of Theorem 11.1)

We are asked to show that

$$Z \perp\!\!\!\perp \{Y(1), Y(0), X\} \mid e(X, Y(1), Y(0)) \tag{11.4}$$

To simplify the notation, we denote $V = \{Y(1), Y(0), X\}$ and $e(V) = P(Z = 1 \mid V)$. Our goal is to show that

$$P(Z = 1 \mid V, e(V)) = P(Z = 1 \mid e(V)).$$

Note that once we know V, we also know $e(V)$ automatically. Thus, the left-hand side can be simplified as

$$\begin{aligned} P(Z = 1 \mid V, e(V)) &= P(Z = 1 \mid V) && \text{(Knowing } e(V) \text{ does not add information)} \\ &= e(V). && \text{(Definition of propensity score)} \end{aligned}$$

As for the right-hand side,

$$\begin{aligned} P(Z = 1 \mid e(V)) &= \mathrm{E}\left\{P(Z = 1 \mid V, e(V)) \mid e(V)\right\} && \text{(Iterated expectation)} \\ &= \mathrm{E}\left\{P(Z = 1 \mid V) \mid e(V)\right\} && \text{(Knowing } e(V) \text{ does not add information)} \\ &= \mathrm{E}\left\{e(V) \mid e(V)\right\} && \text{(Definition of propensity score)} \\ &= e(V). && \text{(Taking out the constant)} \end{aligned}$$

Since both sides are equal to $e(V)$, we have shown (11.4).

Problem 11.2 (Another version of Theorem 11.1)

We are asked to show that

$$Z \perp\!\!\!\perp Y(z) \mid X \implies Z \perp\!\!\!\perp Y(z) \mid e(X)$$

for $z \in \{0, 1\}$.

It suffices to show that

$$P(Z = 1 \mid Y(z), e(X)) = P(Z = 1 \mid e(X)).$$

For the left-hand side,

$$\begin{aligned} &P(Z = 1 \mid Y(z), e(X)) \\ &= \mathrm{E}\left\{P(Z = 1 \mid Y(z), X, e(X)) \mid Y(z), e(X)\right\} && \text{(Iterated expectation)} \\ &= \mathrm{E}\left\{P(Z = 1 \mid Y(z), X) \mid Y(z), e(X)\right\} && \text{(Knowing } e(X) \text{ does not add information)} \\ &= \mathrm{E}\left\{P(Z = 1 \mid X) \mid Y(z), e(X)\right\} && \text{(Ignorability)} \\ &= \mathrm{E}\left\{e(X) \mid Y(z), e(X)\right\} && \text{(Definition of propensity score)} \end{aligned}$$

$$= e(X). \qquad \text{(Taking out the constant)}$$

As for the right-hand side,

$$\begin{aligned}
&P(Z = 1 \mid e(X)) \\
&= \mathrm{E}\left\{P(Z = 1 \mid X, e(X)) \mid e(X)\right\} && \text{(Iterated expectation)} \\
&= \mathrm{E}\left\{P(Z = 1 \mid X) \mid e(X)\right\} && \text{(Knowing } e(X) \text{ does not add information)} \\
&= \mathrm{E}\left\{e(X) \mid e(X)\right\} && \text{(Definition of propensity score)} \\
&= e(X). && \text{(Taking out the constant)}
\end{aligned}$$

Since both sides are equal to $e(X)$, we have completed the proof.

Problem 11.3 (More results on the IPW estimators)

Consider the Horvitz-Thompson (HT) estimator

$$\hat{\tau}^{\mathrm{ht}} = \frac{1}{n}\sum_{i=1}^{n}\frac{Z_i Y_i}{\hat{e}(X_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-Z_i)Y_i}{1-\hat{e}(X_i)}.$$

First, we are asked to show that if we change Y_i to $Y_i + c$ for some constant c, then $\hat{\tau}^{\mathrm{ht}}$ becomes $\hat{\tau}^{\mathrm{ht}} + c(\hat{1}_{\mathrm{T}} - \hat{1}_{\mathrm{C}})$, where

$$\hat{1}_{\mathrm{T}} = \frac{1}{n}\sum_{i=1}^{n}\frac{Z_i}{\hat{e}(X_i)}, \quad \hat{1}_{\mathrm{C}} = \frac{1}{n}\sum_{i=1}^{n}\frac{1-Z_i}{1-\hat{e}(X_i)}.$$

By substituting Y_i with $Y_i + c$, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{Z_i(Y_i+c)}{\hat{e}(X_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-Z_i)(Y_i+c)}{1-\hat{e}(X_i)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{Z_i Y_i}{\hat{e}(X_i)} + \frac{c}{n}\sum_{i=1}^{n}\frac{Z_i}{\hat{e}(X_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-Z_i)Y_i}{1-\hat{e}(X_i)} - \frac{c}{n}\sum_{i=1}^{n}\frac{1-Z_i}{1-\hat{e}(X_i)} \\
&= \hat{\tau}^{\mathrm{ht}} + c(\hat{1}_{\mathrm{T}} - \hat{1}_{\mathrm{C}}).
\end{aligned}$$

This completes the first part.

Next, we are asked to show that

$$\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{Z_i}{e(X_i)}\right\} = 1, \quad \mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1-Z_i}{1-e(X_i)}\right\} = 1.$$

This can be shown by the direct calculation:

$$\begin{aligned}
\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{Z_i}{e(X_i)}\right\} &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left[\frac{Z_i}{e(X_i)}\right] && \text{(Linearity of expectation)} \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left\{\mathrm{E}\left[\frac{Z_i}{e(X_i)}\,\middle|\, X_i\right]\right\} && \text{(Iterated expectation)}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[\frac{\mathrm{E}[Z_i \mid X_i]}{e(X_i)}\right] && \text{(Taking out the constant)} \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[\frac{e(X_i)}{e(X_i)}\right] && \text{(Definition of propensity score)} \\
&= \frac{1}{n}\sum_{i=1}^{n} 1 && \text{(Simplification)} \\
&= 1.
\end{aligned}
$$

The second equality can be shown similarly.

Problem 11.5 (Balancing score and propensity score: more theoretical results)

Let $b(X)$ be a balancing score, i.e., $Z \perp\!\!\!\perp X \mid b(X)$. We are asked to show that $b(X)$ is a balancing score if and only if $e(X)$ is a function of $b(X)$.

We first show the "if" part. Assume that $e(X)$ is a function of $b(X)$; i.e., there exists some function f such that $e(X) = f(b(X))$. We want to show that

$$P(Z = 1 \mid X, b(X)) = P(Z = 1 \mid b(X)).$$

The left-hand side can be simplified as

$$
\begin{aligned}
P(Z = 1 \mid X, b(X)) &= P(Z = 1 \mid X) && \text{(Knowing } b(X) \text{ does not add information)} \\
&= e(X). && \text{(Definition of propensity score)}
\end{aligned}
$$

The right-hand side can be simplified as

$$
\begin{aligned}
&P(Z = 1 \mid b(X)) \\
&= \mathrm{E}\left\{P(Z = 1 \mid X, b(X)) \mid b(X)\right\} && \text{(Iterated expectation)} \\
&= \mathrm{E}\left\{P(Z = 1 \mid X) \mid b(X)\right\} && \text{(Knowing } b(X) \text{ does not add information)} \\
&= \mathrm{E}\left\{e(X) \mid b(X)\right\} && \text{(Definition of propensity score)} \\
&= \mathrm{E}\left\{f(b(X)) \mid b(X)\right\} && \text{(Assumption)} \\
&= f(b(X)) && \text{(Taking out the constant)} \\
&= e(X). && \text{(Assumption)}
\end{aligned}
$$

Thus, we have shown the "if" part.

Next, we show the "only if" part. Assume that $b(X)$ is a balancing score; i.e., $Z \perp\!\!\!\perp X \mid b(X)$. We want to show that

$$e(X) = P(Z = 1 \mid X) \text{ is a function of } b(X).$$

By the definition of balancing score,

$$
\begin{aligned}
&P(Z = 1 \mid X) \\
&= P(Z = 1 \mid X, b(X)) && \text{(Knowing } b(X) \text{ does not add information)} \\
&= P(Z = 1 \mid b(X)). && \text{(Definition of balancing score)}
\end{aligned}
$$

Thus, we have shown the "only if" part.

Problem 11.6 (Some basics of subgroup effects)

Consider a standard observational study with covariates $X = (X_1, X_2)$, where X_1 denotes a binary subgroup indicator and X_2 contains the rest of the covariates. The parameter of interest is the subgroup causal effect

$$\tau(x_1) = \mathrm{E}\left(Y(1) - Y(0) \mid X_1 = x_1\right)$$

First, we are asked to show that

$$\tau(x_1) = \mathrm{E}\left\{\frac{\mathbf{1}\{X_1 = x_1\}ZY}{e(X)} - \frac{\mathbf{1}\{X_1 = x_1\}(1-Z)Y}{1-e(X)}\right\} \Big/ P(X_1 = x_1)$$

under the ignorability assumption. To save space, we only show the first term. Since

$$\mathrm{E}\left[Y(1) \mid X\right] = \mathrm{E}\left[\frac{ZY}{e(X)} \,\middle|\, X\right],$$

we have

$$\begin{aligned}
&\mathrm{E}\left[Y(1) \mid X_1 = x_1\right] \\
&= \mathrm{E}\left\{\mathrm{E}\left[Y(1) \mid X\right] \mid X_1 = x_1\right\} && \text{(Iterated expectation)} \\
&= \mathrm{E}\left\{\mathrm{E}\left[\frac{ZY}{e(X)} \,\middle|\, X\right] \,\middle|\, X_1 = x_1\right\} && \text{(From the previous equation)} \\
&= \mathrm{E}\left\{\frac{ZY}{e(X)} \,\middle|\, X_1 = x_1\right\} && \text{(Iterated expectation)} \\
&= \mathrm{E}\left\{\frac{\mathbf{1}\{X_1 = x_1\}ZY}{e(X)}\right\} \Big/ P(X_1 = x_1). && \text{(Definition of conditional expectation)}
\end{aligned}$$

The second term can be shown similarly. Therefore, we have completed the first part.

To give the HT estimator for $\tau(x_1)$, we can replace the expectations and probabilities by their sample analogues:

$$\begin{aligned}
&\hat{\tau}^{\mathrm{ht}}(x_1) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mathbf{1}\{X_{1i} = x_1\}Z_iY_i}{\hat{e}(X_i)} - \frac{\mathbf{1}\{X_{1i} = x_1\}(1-Z_i)Y_i}{1-\hat{e}(X_i)}\right) \Big/ \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{X_{1i} = x_1\}.
\end{aligned}$$

To give the Hájek estimator for $\tau(x_1)$, we can normalize the weights in the HT estimator:

$$\begin{aligned}
&\hat{\tau}^{\mathrm{hajek}}(x_1) \\
&= \sum_{i=1}^{n}\frac{\mathbf{1}\{X_{1i} = x_1\}Z_iY_i/\hat{e}(X_i)}{\sum_{i=1}^{n}\mathbf{1}\{X_{1i} = x_1\}Z_i/\hat{e}(X_i)} - \sum_{i=1}^{n}\frac{\mathbf{1}\{X_{1i} = x_1\}(1-Z_i)Y_i/(1-\hat{e}(X_i))}{\sum_{i=1}^{n}\mathbf{1}\{X_{1i} = x_1\}(1-Z_i)/(1-\hat{e}(X_i))}.
\end{aligned}$$

Acronyms

HT Horvitz-Thompson. 2, 4