Solutions to Exercises of Chapter 10

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

Problem 10.1 (A simple identity)

Define

$$\begin{aligned}\tau &\equiv \mathrm{E}\left(Y_i(1) - Y_i(0)\right),\\ \tau_{\mathrm{T}} &\equiv \mathrm{E}\left(Y_i(1) - Y_i(0) \mid Z_i = 1\right),\\ \tau_{\mathrm{C}} &\equiv \mathrm{E}\left(Y_i(1) - Y_i(0) \mid Z_i = 0\right),\end{aligned}$$

where the expectations are taken over the joint distribution of $(Y_i(1), Y_i(0), Z_i)$. Later in the derivation, we omit the subscript i for simplicity.

We are asked to show that

$$\tau = P(Z = 1)\tau_{\mathrm{T}} + P(Z = 0)\tau_{\mathrm{C}}.$$

By the law of total expectation,

$$\begin{aligned}\tau &= \mathrm{E}\left(Y(1) - Y(0)\right)\\ &= \mathrm{E}\left(\mathrm{E}\left(Y(1) - Y(0) \mid Z\right)\right)\\ &= P(Z = 1)\,\mathrm{E}\left(Y(1) - Y(0) \mid Z = 1\right) + P(Z = 0)\,\mathrm{E}\left(Y(1) - Y(0) \mid Z = 0\right)\\ &= P(Z = 1)\tau_{\mathrm{T}} + P(Z = 0)\tau_{\mathrm{C}},\end{aligned}$$

which completes the proof.

Problem 10.2 (Nonparametric identification of other causal effects)

Under ignorability, i.e., $Y(z) \perp\!\!\!\perp Z \mid X$ for $z = 0, 1$, we are asked to show that

1. the distributional causal effect

$$P(Y(1) > y) - P(Y(0) > y)$$

is nonparametrically identifiable for all y.
2. the quantile causal effect

$$\mathrm{quantile}_q\left(Y(1)\right) - \mathrm{quantile}_q\left(Y(0)\right)$$

is nonparametrically identifiable for all $q \in (0, 1)$, where $\mathrm{quantile}_q(Y(z)) = \inf\{y : P(Y(z) \le y) \ge q\}$ for $z = 0, 1$.

First, we show that $P(Y(z) > y)$ is nonparametrically identifiable:

$$\begin{aligned}&P(Y(z) > y)\\ &= \mathrm{E}\left(P(Y(z) > y \mid X)\right) && \text{(Iterated Expectation)}\\ &= \mathrm{E}\left(P(Y(z) > y \mid X, Z = z)\right) && \text{(Ignorability)}\end{aligned}$$

$$= \mathrm{E}\left(P(Y > y \mid X, Z = z)\right), \qquad \text{(Consistency)}$$

which suggests that the distributional causal effect is identifiable.

Next, we have

$$\begin{aligned}
&\text{quantile}_q(Y(z)) \\
&= \inf\{y : P(Y(z) \le y) \ge q\} && \text{(Definition of Quantile)} \\
&= \inf\{y : 1 - P(Y(z) > y) \ge q\} && \text{(Complement Rule)} \\
&= \inf\{y : 1 - \mathrm{E}\left(P(Y > y \mid X, Z = z)\right) \ge q\}, && \text{(From the previous part)}
\end{aligned}$$

which suffices to show that the quantile causal effect is identifiable.

Problem 10.3 (Outcome imputation estimator in the fully interacted logistic model)

Assume that a binary outcome follows a logistic model

$$\mathrm{E}\left(Y \mid Z, X\right) = P(Y = 1 \mid Z, X) = \frac{e^{\beta_0 + \beta_z Z + \beta_x X + \beta_{xz}^\top XZ}}{1 + e^{\beta_0 + \beta_z Z + \beta_x X + \beta_{xz}^\top XZ}},$$

We are asked to derive the outcome regression estimator for the average treatment effect (ATE).

By the identification result of the ATE,

$$\begin{aligned}
\tau &= \mathrm{E}\left(Y(1) - Y(0)\right) \\
&= \mathrm{E}\left(\mathrm{E}\left(Y \mid Z = 1, X\right) - \mathrm{E}\left(Y \mid Z = 0, X\right)\right) && \text{(Ignorability \& Consistency)} \\
&= \mathrm{E}\left(\frac{e^{\beta_0 + \beta_z + (\beta_x + \beta_{xz})^\top X}}{1 + e^{\beta_0 + \beta_z + (\beta_x + \beta_{xz})^\top X}} - \frac{e^{\beta_0 + \beta_x^\top X}}{1 + e^{\beta_0 + \beta_x^\top X}}\right), && \text{(From the logistic model)}
\end{aligned}$$

which can be estimated by the sample average

$$\hat{\tau} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{e^{\hat{\beta}_0 + \hat{\beta}_z + (\hat{\beta}_x + \hat{\beta}_{xz})^\top X_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_z + (\hat{\beta}_x + \hat{\beta}_{xz})^\top X_i}} - \frac{e^{\hat{\beta}_0 + \hat{\beta}_x^\top X_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_x^\top X_i}}\right),$$

where $\hat{\beta}_0$, $\hat{\beta}_z$, $\hat{\beta}_x$, $\hat{\beta}_{xz}$ are the maximum likelihood (ML) estimates.

Problem 10.5 (Ignorability versus strong ignorability)

We are asked to give an example where ignorability holds but strong ignorability does not; that is, $Y(z) \perp\!\!\!\perp Z \mid X$ for $z = 0, 1$ but $\{Y(0), Y(1)\} \perp\!\!\!\perp Z \mid X$ does not hold.

Without loss of generality, we omit X in the following example. Let $U \sim \text{Bernoulli}(0.5)$, and $V \sim \text{Bernoulli}(0.5)$ be independent of U. Define $Z = U \oplus V$ (i.e., the XOR operation), and

$$Y(0) = U, \quad Y(1) = V.$$

Then, we have

$$P(Z=1) = P(U \neq V) = 0.5, \quad P(Z=0) = P(U=V) = 0.5,$$

We can verify that

$$P(Y(0)=1 \mid Z=1) = P(U=1 \mid U \neq V) = \frac{P(U=1, V=0)}{P(U \neq V)} = \frac{0.25}{0.5} = 0.5,$$

and similarly,

$$P(Y(0)=1 \mid Z=0) = P(U=1 \mid U=V) = \frac{P(U=1, V=1)}{P(U=V)} = \frac{0.25}{0.5} = 0.5,$$

which implies that $Y(0) \perp\!\!\!\perp Z$. Analogously, we can show that $Y(1) \perp\!\!\!\perp Z$:

$$P(Y(1)=1 \mid Z=1) = P(V=1 \mid U \neq V) = \frac{P(U=0, V=1)}{P(U \neq V)} = \frac{0.25}{0.5} = 0.5,$$

and

$$P(Y(1)=1 \mid Z=0) = P(V=1 \mid U=V) = \frac{P(U=1, V=1)}{P(U=V)} = \frac{0.25}{0.5} = 0.5.$$

However, we have

$$P(Y(0)=1, Y(1)=1 \mid Z=1) = P(U=1, V=1 \mid U \neq V) = 0,$$

and

$$P(Y(0)=1, Y(1)=1 \mid Z=0) = P(U=1, V=1 \mid U=V) = 1,$$

which implies that $\{Y(0), Y(1)\}$ is not independent of Z. Thus, we have constructed an example where ignorability holds but strong ignorability does not.

Acronyms

ATE average treatment effect. 2
ML maximum likelihood. 2