

## MATH 4824C: Causal Inference

### ♠ Solution of Midterm Sample

**Q1.** In randomized trials; i.e.  $(Y_i(1), Y_i(0), X_i) \perp Z_i$ , when  $E[Y(1)] = 0$ , show that  $\text{Var}(ZY) = P(Z = 1)\text{Var}(Y(1))$ .

**Answer**

*Proof.* Observe that

$$\begin{aligned}\text{Var}(ZY) &= E(ZY^2) - (E[ZY])^2 \\ &= E[Z((1 - Z)Y(0) + ZY(1))^2] - (E[ZY(1)])^2 \\ &= E[ZY^2(1)] - E[ZY(1)]^2 \\ &= E(Z) (E[Y^2(1)]) - (EZ)^2 \cdot (EY(1))^2 \\ &= E(Z) (E[Y^2(1)]) \\ &= P(Z = 1) \text{Var}(Y(1)).\end{aligned}$$

□

**Q2.** Show that in observational studies without unmeasured confounders; i.e.  $(Y_i(1), Y_i(0)) \perp Z_i \mid X_i$ ,  $E(Y(1) - X) = E[E(Y - X) \mid Z = 1, X]$ .

**Answer**

*Proof.* Observe that

$$\begin{aligned}E[E(Y - X) \mid Z = 1, X] &= E[E(ZY(1) + (1 - Z)Y(0) \mid Z = 1, X) - X] \\ &= E[E(Y(1) \mid X) - X] \\ &= E(Y(1)) - E(X).\end{aligned}$$

□

**Q3.** Show that in observational studies without unmeasured confounders; i.e.  $(Y_i(1), Y_i(0)) \perp Z_i \mid X_i$ ,  $E(Y(1) - X) = E[Z(Y - X)/e(X)]$  where  $e(X) = P(Z = 1 \mid X)$ .

**Answer**

*Proof.* Observe that

$$\begin{aligned}
& E[Z(Y - X)/e(X)] \\
&= E[Z((1 - Z)Y(0) + ZY(1) - X)/e(X)] \\
&= E[(ZY(1) - ZX)/e(X)] \\
&= E[E((ZY(1) - ZX)/e(X)|X)] \\
&= E[(1/e(X))(e(X)E[Y(1)|X] - e(X)X)] \\
&= E(Y(1)) - E(X).
\end{aligned}$$

□

**Q4.** Show that in observational studies without unmeasured confounders; i.e.  $(Y_i(1), Y_i(0)) \perp Z_i \mid X_i$ , let  $e(x, \beta)$ , be the logistic model of  $Z_i$  on  $X_i$ , and  $\mu_1(x, \alpha_1)$  be linear or logistic model of  $Y_i(1)$  on  $X_i$ . Try to show that

$$\tilde{\mu}_{1,DR} = E \left[ \frac{Z_i \{Y_i - \mu_1(X_i, \alpha_1)\}}{e(X_i, \beta)} + \mu_1(X_i, \alpha_1) \right]$$

is doubly robust, that is, if either  $\mu_1(x, \alpha_1) = \mu_1(x) := E[Y_i|Z_i = 1, X_i = x]$  or  $e(x, \beta) = e(x) := P(Z_i|X_i = x)$ , then  $\tilde{\mu}_{1,DR} = E\{Y_i(1)\}$ .

**Answer**

*Proof.* It is not hard to verify that

$$E[Y_i(1)] = E \left[ \frac{Z_i \{Y_i - \mu_1(X_i)\}}{e(X_i)} + \mu_1(X_i) \right].$$

Then, we observe that

$$\begin{aligned}
& \tilde{\mu}_{1,DR} - E\{Y_i(1)\} \\
&= E \left[ \frac{Z_i \{Y_i - \mu_1(X_i, \alpha_1)\}}{e(X_i, \beta)} + \mu_1(X_i, \alpha_1) \right] - E \left[ \frac{Z_i \{Y_i - \mu_1(X_i)\}}{e(X_i)} + \mu_1(X_i) \right] \\
&= E \left[ \frac{Z_i \{Y_i - \mu_1(X_i, \alpha_1)\}}{e(X_i, \beta)} + \mu_1(X_i, \alpha_1) \right] - E \left[ \frac{Z_i \{Y_i - \mu_1(X_i)\}}{e(X_i, \beta)} + \mu_1(X_i) \right] \\
&\quad + E \left[ \frac{Z_i \{Y_i - \mu_1(X_i)\}}{e(X_i, \beta)} + \mu_1(X_i) \right] - E \left[ \frac{Z_i \{Y_i - \mu_1(X_i)\}}{e(X_i)} + \mu_1(X_i) \right] \\
&= E \left[ \left( 1 - \frac{Z_i}{e(X_i, \beta)} \right) (\mu_1(X_i, \alpha_1) - \mu_1(X_i)) \right] + E \left[ Z_i \{Y_i - \mu_1(X_i)\} \left( \frac{1}{e(X_i, \beta)} - \frac{1}{e(X_i)} \right) \right] \\
&= E \left[ E \left[ \left( 1 - \frac{e(X_i)}{e(X_i, \beta)} \right) (\mu_1(X_i, \alpha_1) - \mu_1(X_i)) \mid X_i \right] \right] \\
&\quad + E \left[ E \left[ e(X_i) \{ \mu_1(X_i) - \mu_1(X_i) \} \left( \frac{1}{e(X_i, \beta)} - \frac{1}{e(X_i)} \right) \mid X_i \right] \right].
\end{aligned}$$

The conclusion can be directly induced by the resulting form.

□

**Q5.** In observational studies with binary response and binary unmeasured confounders  $U_i$ ; i.e.  $(Y_i(1), Y_i(0)) \perp Z_i \mid (X_i, U_i)$ , if  $RR_{ZY}^{obs} = 1.2$ , calculate the  $E$ -value and interpret it.

**Answer**

*Proof.* E-value is defined as

$$RR_{ZY}^{obs} + \sqrt{RR_{ZY}^{obs} (RR_{ZY}^{obs} - 1)}.$$

In this case, it equals

$$1.2 + \sqrt{1.2 \times 0.2}.$$

Interpretation: see page 20 of Chapter 5 of the lecture notes.

□