CAUSAL INFERENCE Part III / Michaelmas 2023
Example Sheet 3 (of 3)

1. In lectures, we discussed design and modelling bias of causal estimators. How large are these biases for the regression adjustment estimators for Bernoulli trials?

2. Suppose the treatment A is binary. Let $\pi(\boldsymbol{x}) = \mathbb{P}(A = 1 \mid \boldsymbol{X} = \boldsymbol{x})$ be the propensity score. Under the no unmeasured confounders assumption, prove that

$$A \perp\!\!\!\perp Y(a) \mid \pi(\boldsymbol{X}), \text{ for } a = 0, 1.$$

A function of the covariates $b(\boldsymbol{X})$ is called a *balancing score* if it satisfies $A \perp\!\!\!\perp \boldsymbol{X} \mid b(\boldsymbol{X})$. Show that $\pi(\boldsymbol{X})$ is a balancing score and can be written as a function of any other balancing score $b(\boldsymbol{X})$.

3. Consider the matched pair design of observational studies in which observation i is matched to observation $i + n_1$, $i = 1, \ldots, n_1$. Suppose the data are iid and there are no unmeasured confounders. Let $\boldsymbol{C}_i = (\boldsymbol{X}_i, Y_i(0), Y_i(1))$ and

$$M = \{\boldsymbol{a}_{[2n_1]} \in \{0,1\}^{2n_1} \mid a_i + a_{i+n_1} = 1, \forall i \in [n_1]\}$$

be all the treatment assignments such that exactly one observation receives the treatment in each matched pair. Show that if $\pi(\boldsymbol{X}_i) = \pi(\boldsymbol{X}_{i+n_1})$ for all $i \in [n_1]$, matching recreates a pairwise randomised experiment in the sense that

$$\mathbb{P}\left(\boldsymbol{A}_{[2n_1]} = \boldsymbol{a} \,\middle|\, \boldsymbol{C}_{[2n_1]}, \boldsymbol{A}_{[2n_1]} \in M\right) = \begin{cases} 2^{-n_1}, & \text{if } \boldsymbol{a} \in M, \\ 0, & \text{otherwise.} \end{cases}$$

4. Find the influence function for the regression estimator $\hat{\beta}_1$ in the analysis of Bernoulli trials. Veryify that it has mean 0.

5. Let the parameter β be defined as the unique solution to

$$\mathbb{E}_{\mathbb{P}}\{m(\beta; V)\} = 0.$$

Show that the influence function of β is given by

$$\phi_{\mathbb{P}}(V) = -\mathbb{E}_{\mathbb{P}}\Big[\Big\{\frac{\partial}{\partial \beta} m(\beta, V)\Big\}^{-1}\Big] m(\beta, V).$$

In your derivation, you may assume suitable differentiability of the function m and any differentiation can be interchanged with expectation.

6. Consider an i.i.d. sample $\boldsymbol{V}_i = (X_i, A_i, Y_i)$, $i = 1, \ldots, n$ with binary A_i. Suppose X is discrete. For any a and x, let

$$\hat{\pi}_a(x) = \frac{\sum_{i=1}^n 1_{\{A_i = a, X_i = x\}}}{\sum_{i=1}^n 1_{\{X_i = x\}}},$$

and

$$\hat{\mu}_a(x) = \frac{\sum_{i=1}^n 1_{\{A_i = a, X_i = x\}} Y_i}{\sum_{i=1}^n 1_{\{A_i = a, X_i = x\}}}$$

be nonparametric estimators of $\pi_a(x) = \mathbb{P}(A = 1 \mid X = x)$ and $\mu_a(x) = \mathbb{E}[Y \mid A = a, X = x]$ (suppose the denominators in the above expressions are positive). Show that the outcome regression estimator

$$\hat{\beta}_{\text{OR}} = \frac{1}{n} \sum_{i=1}^n \hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)$$

is equal to the inverse probability weighted estimator

$$\hat{\beta}_{\mathrm{IPW}} = \frac{1}{n}\sum_{i=1}^{n}\Big[\frac{A_i}{\hat{\pi}(X_i)} - \frac{1-A_i}{1-\hat{\pi}(X_i)}\Big]Y_i.$$

Furthremore, show that for any a and function $\mu(x)$,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1_{\{A_i=a\}}}{\hat{\pi}_a(X_i)}\mu(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu(X_i).$$

7. Continuing from the last question, let

$$m_a(\boldsymbol{V};\mu_a,\pi_a) = \frac{1_{\{A=a\}}}{\pi_a(X)}\big(Y-\mu_a(X)\big)+\mu_a(X),\ a=0,1.$$

   (a) Under the positivity assumption $\pi_a(x)>0,\forall x$, show that for any functions $\tilde{\mu}_a(x)$ and $\tilde{\pi}_a(x)$,

$$\beta_a := \mathbb{E}[\mu_a(X)] = \mathbb{E}\left[\frac{1_{\{A=a\}}}{\pi_a(X)}Y\right] = \mathbb{E}[m_a(D;\mu_a,\pi_a)] = \mathbb{E}[m_a(D;\mu_a,\tilde{\pi}_a)] = \mathbb{E}[m_a(D;\tilde{\mu}_a,\pi_a)].$$

   (b) Consider the estimator

$$\hat{\beta}_{a,\mathrm{DR}} = \frac{1}{n}\sum_{i=1}^{n}m_a(\boldsymbol{V}_i;\hat{\mu}_a,\hat{\pi}_a),$$

   where $\hat{\mu}_a(x)$ and $\hat{\pi}_a(x)$ are obtained by fitting some parametric models. Outline an argument that shows $\hat{\beta}_{a,\mathrm{DR}}$ is doubly robust in the sense that $\hat{\beta}_{a,\mathrm{DR}}$ consistently estimates β_a if at least one of the parametric models for $\hat{\mu}_a(x)$ and $\hat{\pi}_a(x)$ are correctly specified.

8. Consider the instrumental variables estimator

$$\hat{\beta}_g = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i-\bar{Y})g(Z_i)}{\frac{1}{n}\sum_{i=1}^{n}(A_i-\bar{A})g(Z_i)}.$$

   (a) Veryify that if Z is binary,

$$\frac{\mathrm{Cov}(Z,Y)}{\mathrm{Cov}(Z,A)} = \frac{\mathbb{E}[Y\mid Z=1]-\mathbb{E}[Y\mid Z=0]}{\mathbb{E}[A\mid Z=1]-\mathbb{E}[A\mid Z=0]}.$$

   (b) Prove the asymptotic normality of $\hat{\beta}_g$ by showing the influence function of $\hat{\beta}_g$ is given by

$$\psi_g(Z,A,Y) = \frac{[\{Y-\mathbb{E}(Y)\}-\beta\{A-\mathbb{E}(A)\}][g(Z)-\mathbb{E}\{g(Z)\}]}{\mathrm{Cov}(A,g(Z))}.$$

   (c) Derive the above expression again by applying the calculus of influence functions to the functional $\beta(\mathbb{P}) = \mathrm{Cov}(Y,g(Z))/\mathrm{Cov}(A,g(Z))$.

9. Consider a sensitivity analysis that specifies $\delta_a(x) = \mathbb{E}[Y(a)\mid A=1,X=x]-\mathbb{E}[Y(a)\mid A=0,X=x], a=0,1$. Show that the design bias for estimating the average treatment effect is given by

$$\mathbb{E}\{\mathbb{E}[Y\mid A=1,X]\}-\mathbb{E}\{\mathbb{E}[Y\mid A=0,X]\}-\mathbb{E}[Y(1)-Y(0)] = \mathbb{E}\left[(1-\pi(X))\delta_1(X)+\pi(X)\delta_0(X)\right].$$

   Use this to suggest an outcome regression estimator and a doubly robust estimator for the average treatment effect.

10. Consider the causal diagram below where U is unobserved. Suppose the negative control outcome W has the same confounding bias as Y in the following sense:

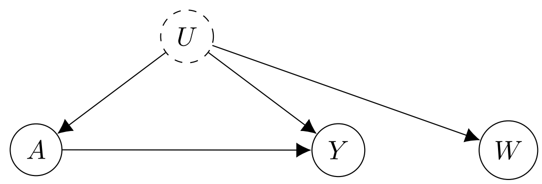
$$\mathbb{E}[Y(0) \mid A=1] - \mathbb{E}[Y(0) \mid A=0] = \mathbb{E}[W(0) \mid A=1] - \mathbb{E}[W(0) \mid A=0].$$

Show that the so-called *parallel trend* assumption

$$\mathbb{E}[Y(0) - W \mid A=1] = \mathbb{E}[Y(0) - W \mid A=0]$$

is satisfied, and use it to show that the average treatment effect on the treated is identified by the *difference-in-differences* estimator:

$$\mathbb{E}[Y(1) - Y(0) \mid A=1] = \mathbb{E}[Y - W \mid A=1] - \mathbb{E}[Y - W \mid A=0].$$



11. Continuing from Question 3, suppose there exists an unmeasured confounder $U \in [0,1]$ so that $A \perp\!\!\!\perp \{Y(0), Y(1)\} \mid X, U$. Let $\pi_i = \mathbb{P}(A_i = 1 \mid X_i, U_i)$. Show that the logistic regression model

$$\mathbb{P}(A=1 \mid X, U) = \operatorname{expit}(g(X) + \gamma U), \ \ 0 \le \gamma \le \log \Gamma,$$

where $g(\cdot)$ is an arbitrary function, $\operatorname{expit}(\eta) = e^{\eta}/(1+e^{\eta})$ and $\Gamma \ge 1$ is a constant, implies Rosenbaum's sensitivity model

$$\frac{1}{\Gamma} \le \frac{\pi_i/(1-\pi_i)}{\pi_{n_1+i}/(1-\pi_{n_1+i})} \le \Gamma, \ \ \forall i \in [n_1].$$

For the sign test (see ES1 Q8b), derive a "upper bounding" asymptotic randomization p-value for this model. This p-value should be valid for all randomization distributions in the above sensitivity model.