1. (Exercise 6.3) How much are the design bias and modelling bias of the regression adjustment estimators for Bernoulli trials?

2. (Exercise 6.6) 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

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

Furthermore, show that for any $b(\boldsymbol{x})$ that satisfies $A \perp\!\!\!\perp Y \mid b(\boldsymbol{X})$, $\pi(\boldsymbol{x})$ can be written as a function of $b(\boldsymbol{x})$.

3. (Exercise 6.9) Consider the matched pair design of observational studies, so observation i is matched to observation $i + n_1$, $i = 1, \dots, n_1$. 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. (Exercise 6.13) Consider the signed score statistic defined in the lectures. Derive the randomisation test based on the randomisation distribution of $T(\boldsymbol{A}_{[2n_1]}, \boldsymbol{Y}_{[2n_1]}(0))$ and show that, under $H_0 : Y_i(1) - Y_i(0) = \beta, i \in [n_!]$ and conditioning on $\boldsymbol{C}_{[2n_1]}$,

$$T\big(\boldsymbol{A}_{[2n_1]}, \boldsymbol{Y}_{[2n_1]}(0)\big) \mid \boldsymbol{A}_{[2n_1]} \in M \overset{d}{=} \sum_{i=1}^{n_1} S_i \psi\Big(\frac{\text{rank}(|Y_i(0) - Y_{n_1+i}(0)|)}{n_1 + 1}\Big), \tag{1}$$

where $S_i = (A_i - A_{i+n_1}) \cdot \text{sgn}(Y_i(0) - Y_{i+n_1}(0)) \sim \text{Bernoulli}(1/2)$. Justify this test using the symmetry of $D_i - \beta$ under H_0.

5. (Exercise 7.5) Derive the influence function for the regression estimator $\hat{\beta}_1$ in analysing Bernoulli trials. Veryify that it has mean 0.

6. (Exercise 7.9) Suppose $\boldsymbol{X}$ is discrete. Let $\hat{\pi}_a(\boldsymbol{x}) = [\sum_{i=1}^n I(A_i = a, \boldsymbol{X}_i = \boldsymbol{x})]/[\sum_{i=1}^n I(\boldsymbol{X}_i = \boldsymbol{x})] > 0$, $\forall a, \boldsymbol{x}$ and $\hat{\mu}_a(\boldsymbol{x}) = [\sum_{i=1}^n I(A_i = a, \boldsymbol{X}_i = \boldsymbol{x})Y_i]/[\sum_{i=1}^n I(A_i = a, \boldsymbol{X}_i = \boldsymbol{x})]$ be nonparametric estimators of $\pi_a(\boldsymbol{x}) = \mathbb{P}(A = 1 \mid \boldsymbol{X} = \boldsymbol{x})$ and $\mu_a(\boldsymbol{x}) = \mathbb{E}[Y \mid A = a, \boldsymbol{X} = \boldsymbol{x}]$. Show that

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

is equal to

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

Furthremore, show that for any function $\mu(\boldsymbol{x})$,

$$\frac{1}{n} \sum_{i=1}^n \frac{I(A_i = a)}{\hat{\pi}_a(\boldsymbol{X}_i)} \mu(\boldsymbol{X}_i) = \frac{1}{n} \sum_{i=1}^n \mu(\boldsymbol{X}_i), \ a = 0, 1.$$

7. (Exercises 7.14, 7.18, 7.19) Consider the definitions of $\pi(\boldsymbol{x})$ and $\mu(\boldsymbol{x})$ in the last question. Suppose $\boldsymbol{D}_i = (\boldsymbol{X}_i, A_i, Y_i)$ are iid and let

$$m_a(\boldsymbol{D}_i; \mu_a, \pi_a) = \frac{I(A_i = a)}{\pi_a(\boldsymbol{X}_i)}\big(Y_i - \mu_a(\boldsymbol{X}_i)\big) + \mu_a(X_i),\ a = 0, 1.$$

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

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

   (b) Consider the estimator

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

   where $\hat{\mu}_a(\boldsymbol{x})$ and $\hat{\pi}_a(\boldsymbol{x})$ 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 $\hat{\mu}_a(\boldsymbol{x})$ and $\hat{\pi}_a(\boldsymbol{x})$ are correctly specified.

8. (Exercise 8.3) Consider the setting in Question 3. Suppose there exists an unmeasured confounder $U \in [0, 1]$ so that $A \perp\!\!\!\perp \{Y(0), Y(1)\} \mid \boldsymbol{X}, U$. Let $\pi_i = \mathbb{P}(A_i = 1 \mid \boldsymbol{X}_i, U_i)$. Show that the logistic regression model

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

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

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

9. (Exercise 8.5) For the sign test (corresponding to the choice $\psi(t) \equiv 1$ for the signed score statistics), derive an asymptotic p-value for Rosenbaum's sensitivity analysis based on a central limit theorem for the bounding variable in the lecture notes.

10. (Exercises 8.6, 8.7) Consider a sensitivity analysis that specifies $\delta_a(\boldsymbol{x}) = \mathbb{E}[Y(a) \mid A = 1, \boldsymbol{X} = \boldsymbol{x}] - \mathbb{E}[Y(a) \mid A = 0, \boldsymbol{X} = \boldsymbol{x}], a = 0, 1$. Show that the design bias for estimating the average treatment effect is given by

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

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

11. (Exercise 9.1) Consider the causal diagram below. Suppose the negative control outcome W has the same confounding bias as Y in the following sense:

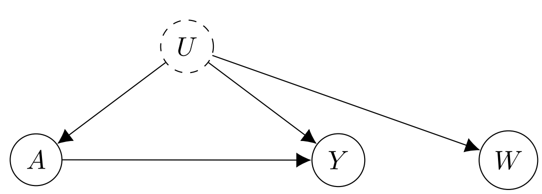
$$\mathbb{E}[Y(0) \mid A = 1] - \mathbb{E}[Y(0) \mid A = 0] = \mathbb{E}[W(0) \mid A = 1] - \mathbb{E}[W(0) \mid A = 0].$$

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 so-called difference-in-differences estimator:

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



12. (Exercises 9.6, 9.7, 9.10) Consider the method of moments IV estimator

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

(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

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

(c) Suppose we estimate the optimal instrument $g^*(\boldsymbol{Z}) = \mathbb{E}[A \mid \boldsymbol{Z}]$ by fitting a linear model using least squares. Show that $\hat{\beta}_{\hat{g}}$ is the two-stage least squares estimator described in the lectures.

13. To learn how to apply the theory and methods of causal inference you learned, the best way is to read some applied articles. Sign up for an applied article at `http://bit.ly/33i883v` and present it in the 4th example class.