DS-GA 3001.001 Applied Statistics: Homework #4

Due on Thursday, November 21, 2024

Please hand in your homework via Gradescope (entry code: DKYKGY) before 11:59 PM.

1. Revisit the example of bivariate Gaussian location model we covered in class:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \cdots, \begin{bmatrix} x_n \\ y_n \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \theta_0 \\ \eta_0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right),$$

where $\rho \in [-1, 1]$ is known.

(a) Recall that the estimating equation based on the score for θ_0 is

$$\frac{1}{n} \sum_{i=1}^{n} \left[x_i - \widehat{\theta} - \rho(y_i - \widehat{\eta}) \right] = 0.$$

If $\widehat{\eta} = \eta_0$ is the true nuisance, from the above equation, determine the probability distribution of $\widehat{\theta} - \theta_0$ which only depends on (n, ρ).

(b) Repeat (a) if $\widehat{\eta} = \eta_0 + \varepsilon$ with a fixed constant ε. Your answer should depend on (n, ρ, ε).

(c) Now consider the efficient score equation

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - \widehat{\theta}) = 0.$$

Write out the probability distribution of $\widehat{\theta} - \theta_0$. How does $\mathbb{E}[(\widehat{\theta} - \theta_0)^2]$ compare with (a) and (b)?

2. In the estimation of ATE, in class we modeled the mean outcomes for each group:

$$\mu_0(X) = \mathbb{E}[Y \mid X, W = 0], \quad \mu_1(X) = \mathbb{E}[Y \mid X, W = 1].$$

Another modeling is to model a single mean outcome $m(X) = \mathbb{E}[Y \mid X]$ and consider the following estimating function

$$f_{(m,e,\tau)}(W, X, Y) = (Y - m(X) - (W - e(X))\tau)(W - e(X)),$$

where $e(X) = \mathbb{P}(W = 1 \mid X)$ is the propensity score.

(a) Find the expression of $m(X)$ in terms of $(\mu_0(X), \mu_1(X), e(X))$.

(b) Assuming that $\mu_1(x) = \mu_0(x) + \tau$ for all x, show that $f_{(m,e,\tau)}(W, X, Y)$ is a valid estimating function, i.e.

$$\mathbb{E}[f_{(m,e,\tau)}(W, X, Y)] = 0.$$

(c) Show that $f_{(m,e,\tau)}(W, X, Y)$ is Neyman orthogonal with respect to (m, e), i.e.

$$\mathbb{E}[\nabla_m f_{(m,e,\tau)}(W, X, Y)] = 0,$$
$$\mathbb{E}[\nabla_e f_{(m,e,\tau)}(W, X, Y)] = 0.$$

(d) Show that $f_{(m,e,\tau)}(W, X, Y)$ is *not* doubly robust, by arguing that in general

$$\mathbb{E}[f_{(m,\widehat{e},\tau)}(W, X, Y)] \neq 0.$$

3. Consider the same setting for the AIPW estimator in class, but now we aim to estimate the average treatment effect on the treated (ATTE): $\tau^{\mathrm{ATTE}} = \mathbb{E}[\mu_1(X) - \mu_0(X) \mid W = 1]$. Consider the following estimating function:

$$f_{(\mu_0,e,\tau^{\mathrm{ATTE}})}(W, X, Y) = \frac{W(Y - \mu_0(X) - \tau^{\mathrm{ATTE}})}{m} - \frac{e(X)(1 - W)(Y - \mu_0(X))}{m(1 - e(X))},$$

where $e(x) = \mathbb{P}(W = 1 \mid X = x)$ is the propensity score, and $m = \mathbb{P}(W = 1)$ is the marginal probability of treatment. For simplicity we assume that m is known.

(a) Let $p(x)$ be the pmf of $X = x$. Using the Bayes rule, show that

$$\mathbb{P}(X = x \mid W = 1) = \frac{p(x)e(x)}{m}.$$

(b) Use (a) to prove the following identity:

$$\tau^{\mathrm{ATTE}} = \mathbb{E}\left[\frac{e(X)}{m}(\mu_1(X) - \mu_0(X))\right].$$

(c) Show that $f_{(\mu_0,e,\tau^{\mathrm{ATTE}})}(W, X, Y)$ is a valid estimating function, i.e.

$$\mathbb{E}[f_{(\mu_0,e,\tau^{\mathrm{ATTE}})}(W, X, Y)] = 0.$$

(d) (*Bonus 5 points*) Show that $f_{(\mu_0,e,\tau^{\mathrm{ATTE}})}(W, X, Y)$ is doubly robust, i.e. for any $(\widehat{\mu}_0(x), \widehat{e}(x))$,

$$\mathbb{E}[f_{(\widehat{\mu}_0,e,\tau^{\mathrm{ATTE}})}(W, X, Y)] = 0,$$
$$\mathbb{E}[f_{(\mu_0,\widehat{e},\tau^{\mathrm{ATTE}})}(W, X, Y)] = 0.$$

4. Coding I: we will implement Stein's semiparametric estimator for the symmetric location model $y_1, \cdots, y_n \sim f(y - \theta_0)$, where in our experiment $f(y) = e^{-|y|}/2$ is the Laplace density. We will experiment on three estimators of θ_0:

- the sample mean of $(y_1, \cdots, y_n)$;
- the MLE with the knowledge of f - you should derive the form of the MLE here and find it to be a very simple statistic of $(y_1, \cdots, y_n)$;
- Stein's semiparametric estimator without the knowledge of f.

Based on inline instructions, fill in the missing codes in `https://tinyurl.com/5zjf4bzd`. Be sure to submit a pdf with your codes, outputs, and colab link.

5. Coding II: we will compare the IPW and AIPW estimators on a synthetic dataset. Based on inline instructions, fill in the missing codes in `https://tinyurl.com/y22fams3`. Be sure to submit a pdf with your codes, outputs, and colab link.