Day 1, Practical 2

Helene Charlotte Wiese Rytgaard

May 11, 2023

Recall that the estimating equation (ee) estimator is defined by the following procedure:

1. Estimate nuisance parameters $f(a,x) = \mathbb{E}[Y \mid A = a, X = x]$, $\pi(a \mid x) = \mathbb{E}[A \mid X = x]$ and the average over the distribution P of O.

2. Plug in to estimate the ATE:

$$\begin{aligned}\hat{\psi}_n^{\text{ee}} &= \tilde{\Psi}_{\text{ee}}(\hat{f}_n, \hat{\pi}_n, \hat{P}_n) \\ &= \frac{1}{n}\left\{\left(\frac{A_i}{\hat{\pi}_n(1 \mid X_i)} - \frac{1-A_i}{\hat{\pi}_n(0 \mid X_i)}\right)\big(Y_i - \hat{f}_n(A_i, X_i)\big) + \hat{f}_n(1, X_i) - \hat{f}_n(0, X_i)\right\}. \end{aligned} \tag{1}$$

As we have seen, this estimator uses the representation for the target parameter:

$$\begin{aligned}&\tilde{\Psi}_{\text{ee}}(f, \pi, p) \\ &\quad = \mathbb{E}_P\left[\left(\frac{A}{\pi(A \mid X)} - \frac{1-A}{\pi(A \mid X)}\right)\big(Y - f(A, X)\big) + f(1, X) - f(0, X)\right],\end{aligned}$$

involving really an average over all but the last terms of the efficient influence curve:

$$\begin{aligned}&\phi^*(P)(O) = \tilde{\phi}^*(f, \pi)(O) \\ &\quad = \left(\frac{A}{\pi(A \mid X)} - \frac{1-A}{\pi(A \mid X)}\right)\big(Y - f(A, X)\big) + f(1, X) - f(0, X) - \Psi(P).\end{aligned}$$

Particularly, $\hat{\psi}_n^{\text{ee}}$ solves by construction the efficient influence equation:

$$\begin{aligned}\mathbb{P}_n\tilde{\phi}^*(\hat{f}_n, \hat{\pi}_n) &= \frac{1}{n}\sum_{i=1}^n \tilde{\phi}^*(\hat{f}_n, \hat{\pi}_n) \\ &= \frac{1}{n}\left\{\left(\frac{A_i}{\hat{\pi}_n(1 \mid X_i)} - \frac{1-A_i}{\hat{\pi}_n(0 \mid X_i)}\right)\big(Y_i - \hat{f}_n(A_i, X_i)\big) + \hat{f}_n(1, X_i) - \hat{f}_n(0, X_i)\right\} - \hat{\psi}_n^{\text{ee}} \\ &= 0.\end{aligned}$$

A TMLE estimator also solves the efficient influence curve equation, just in a different way. Particularly, the two estimators have the exact same asymptotic properties (but may, however, still differ in finite samples). Recall the following decomposition in analyzing the large-sample properties of an estimator:

$$\hat{\psi}_n^{\text{ee}} - \Psi(P_0) = \mathbb{P}_n\phi^*(P_0) + o_P(n^{-1/2}) + R(\hat{P}_n, P_0) - \underbrace{\mathbb{P}_n\phi^*(\hat{P}_n)}_{=0};$$

when $R(\hat{P}_n, P_0) = o_P(n^{-1/2})$, then $\Psi(\hat{P}_n) \overset{as}{\sim} N(\Psi(P_0), P_0\phi^*(P_0)^2/n)$, and the variance of the estimator can be estimated by

$$\hat{\sigma}_n^2 = \mathbb{P}_n\{\tilde{\phi}^*(\hat{f}_n, \hat{\pi}_n)\}^2/n = \frac{1}{n^2}\sum_{i=1}^n\{\tilde{\phi}^*(\hat{f}_n, \hat{\pi}_n)(O)\}^2 \tag{2}$$

1 Simulate data

We will work with the simulation function defined in the first practical.

Task 1: Use the simulation function from the first practicals from day 1 (Task 1) to draw a random dataset with sample size $n = 1000$.

```
set.seed(15)
head(sim.data <- sim.fun(1000))
```

2 Implement the estimating equation estimator

Task 2: Implement the estimating equation estimator, as outlined below:

1. Fit the models below for the outcome regression f and the propensity score π.
2. Use `fit.f` to predict the conditional expectations $\mathbb{E}_P[Y \mid A, X]$ and $\mathbb{E}_P[Y \mid A = a, X]$. Add these as columns to the dataset.
3. Use `fit.pi` to estimate the propensity score $\pi(a \mid X) = P(A = a \mid X)$. Add this as a column to the simulated dataset.
4. Implement $\hat{\psi}_n^{\text{ee}}$ based on Equation (1).
5. Implement the variance estimator based on Equation (2).

3 Compare with the TMLE estimator

Task 3. Load the `tmle` package and use the `tmle()` function to get the TMLE estimate and variance using the same models as in **Task 2** using the code below. Check that you get about the same.

```
library(tmle)
tmle.fit <- tmle(Y=sim.data$Y, A=sim.data$A,
         cbind(X1=sim.data$X1,
               X2=sim.data$X2,X3=sim.data$X3),
         gform=A~X1+X2+X3, ## treatment model
         Qform=Y~A+X1+X2+X3, ## outcome model
         family="binomial",
         cvQinit=FALSE)
#-- get the ATE estimate:
tmle.fit$estimates$ATE$psi
#-- get the variance estimate:
tmle.fit$estimates$ATE$var
```

Note: As mentioned, the two estimators are supposed to have the same asymptotic properties, but may have different finite-sample properties.