Notation

Helene Charlotte Wiese Rytgaard

September 26, 2021

1 Operators on functions of the observed data

- ▷ For a function $h : \mathcal{O} \to \mathbb{R}$ and distribution P

$$Ph = \mathbb{E}_P[h(O)] = \int h dP = \int_{\mathcal{O}} h(o) dP(o)$$

where $\mathcal{O}$ is the sample space of the observed data O, e.g., $\mathcal{O} = \mathbb{R}^d \times \{0,1\} \times \{0,1\}$ is the sample space of $O = (X, A, Y)$.

- ▷ For the empirical measure $\mathbb{P}_n$ of the sample $O_1, \dots, O_n$:

$$\mathbb{P}_n h = \int h d\mathbb{P}_n = \frac{1}{n} \sum_{i=1}^{n} h(O_i),$$

where the right-hand-side is really just the empirical average.

2 Target parameter

$\Psi : \mathcal{M} \to \mathbb{R}$ (we focus on real-valued).
$\mathcal{M}$ is the statistical model assumed to include P_0, the distribution of the observed data O.
The true value of the target parameter is $\psi_0 = \Psi(P_0)$.
We assume that $\mathcal{M}$ is a nonparametric model.

3 Nuisance parameters

$f(a,x) = \mathbb{E}_P[Y \mid A = a, X = x]$.
$f_0(a,x) = \mathbb{E}_{P_0}[Y \mid A = a, X = x]$.
$\pi(1 \mid x) = \mathbb{E}_P[A \mid X = x]$.
$\pi_0(1 \mid x) = \mathbb{E}_{P_0}[A \mid X = x]$.
$\mu_X(x)$ is the marginal density of X. We also write $d\mu_X = \mu_X \nu_X$ for the appropriate dominating measure ν_X.
$\mu_Y(y \mid a, x)$ is the conditional density of $Y \in \{0,1\}$ given $A = a$ and $X = x$.

4 $o_P(1)$ and $O_P(1)$

$o_P(1)$ denotes a sequence which is converges to zero in probability.
$X_n = o_P(1/r_n)$ if $r_n X_n = o_P(1)$.
$O_P(1)$ denotes a sequence which is bounded in probability.
$X_n = O_P(1/r_n)$ if $r_n X_n = O_P(1)$.
If a sequence is $o_P(1)$ then it is $O_P(1)$.
Also useful to keep in mind that,

$$o_P(1) + o_P(1) = o_P(1), \quad O_P(1) + o_P(1) = O_P(1), \quad \text{and,} \;\; O_P(1) o_P(1) = o_P(1).$$

5 Asymptotic linearity

An estimator $\hat{\psi}_n$ is $\sqrt{n}$-consistent and asymptotically linear with influence function $\phi(P_0)(O)$ if,

$$\sqrt{n}\big(\hat{\psi}_n - \psi_0\big) = \sqrt{n}\,\mathbb{P}_n \phi(P_0) + o_P(1) = \frac{1}{n}\sum_{i=1}^{n} \phi(P_0)(O_i) + o_P(1).$$

influence function ϕ has zero mean and finite variance, i.e., $\mathbb{E}_{P_0}[\phi(P_0)(O)] = 0$ and $\mathbb{E}_{P_0}[\{\phi(P_0)(O)\}^2] < \infty$ (note that this is the same as writing $P_0\phi(P_0) = 0$ and $P_0\{\phi(P_0)\}^2 < \infty$).

6 "Roadmap"

1. Data is a random variable O with a probability distribution P_0.
2. P_0 belongs to a statistical model $\mathcal{M}$.
3. Our target is a parameter $\Psi : \mathcal{M} \to \mathbb{R}$. We wish to estimate $\psi_0 = \Psi(P_0)$.
4. Construct estimator $\hat{P}_n$ for (relevant part of) P_0 and estimate the target parameter by $\hat{\psi}_n = \Psi(\hat{P}_n)$ (substitution estimation).
5. Quantifying uncertainty for the estimator $\hat{\psi}_n = \Psi(\hat{P}_n)$ in terms of asymptotic normal distribution.