Potential Outcome Framework

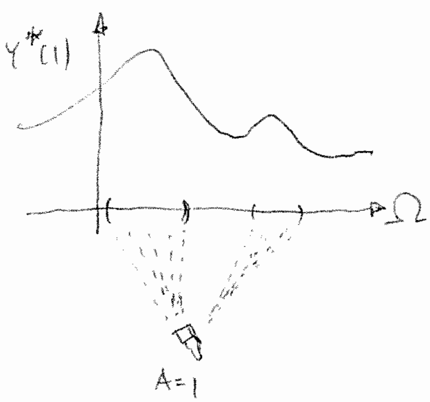
① Observe (Y, A). What is the "causal effect of A on Y?

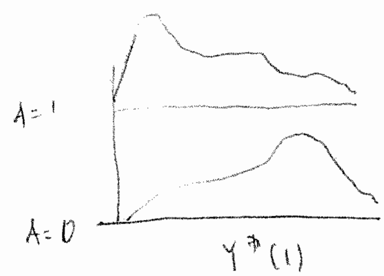
Postulate potential outcomes $Y^*(0), Y^*(1), \ldots$ for $A=0, A=1, \ldots$ (SUTVA)

Link to the data: $Y = AY^*(1) + (1-A)Y^*(0)$ ("consistency")

ATE: $\mathbb{E}Y^*(1) - \mathbb{E}Y^*(0)$

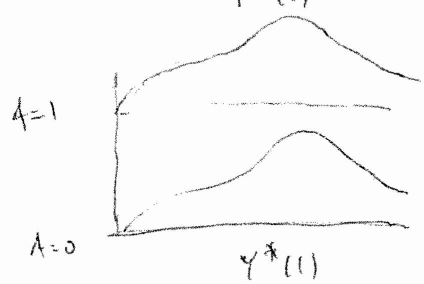
② Estimation - confounding - formulation as a missing data problem





confounding

$\mathbb{E}(Y|A=1) \neq \mathbb{E}Y^*(1)$

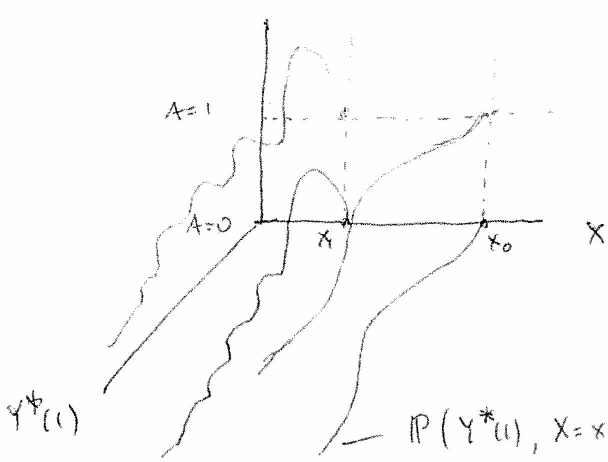


randomization. $A \perp Y^*(1)$

ATE = $\mathbb{E}(Y|A=1) - \mathbb{E}(Y|A=0)$

association = causation

~~Require to ask~~ In absence of randomization, require $A \perp (Y^*(0), Y^*(1)) \mid X$



No unmeasured confounders

$\mathbb{P}(Y^*(1), X=x_0, A=0) \propto \mathbb{P}(Y^*(1), X=x_0, A=1)$

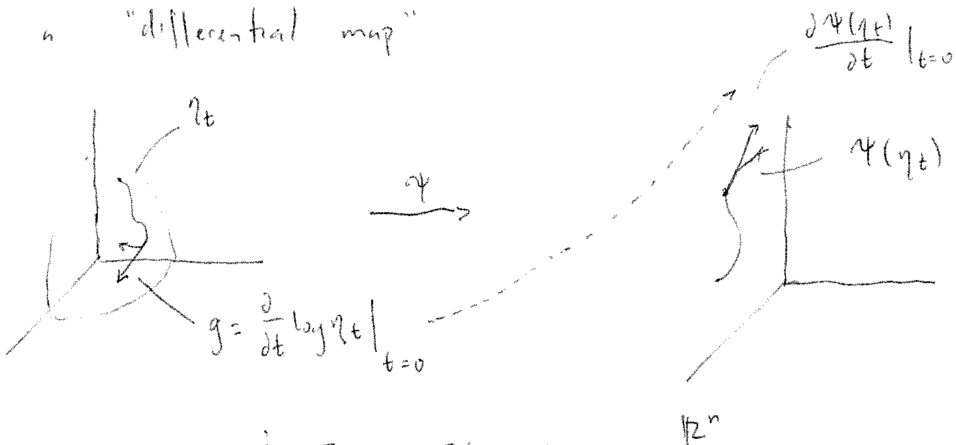
Semiparametrics

① Problem is to estimate $\Psi = \mathbb{E}Y^*(1) - \mathbb{E}Y^*(0)$ on observing (Y, A, X)

Consider RAL estimators $\sqrt{n}(\hat{\delta}_n - \delta) = n^{-1/2} \sum_{j=1}^{n} \dot{\Psi}(Y_j, A_j, X_j) + o_{p_n}(1)$

Geometric interpretation of influence function $\dot{\Psi}$

Given probability model $\eta \in H$ and a functional $\Psi : H \to \mathbb{R}^n$, $\dot{\Psi}$ is like a "differential map"



$$\dot{\Psi} : T(\eta) \to T(\Psi(\eta))$$

$$\dot{\Psi} g = (\dot{\Psi}, g) = \frac{\partial \Psi(\eta_t)}{\partial t}$$

(semiparametric set-up: $\eta = (\theta, v)$, $\Psi(\eta) = \theta$)

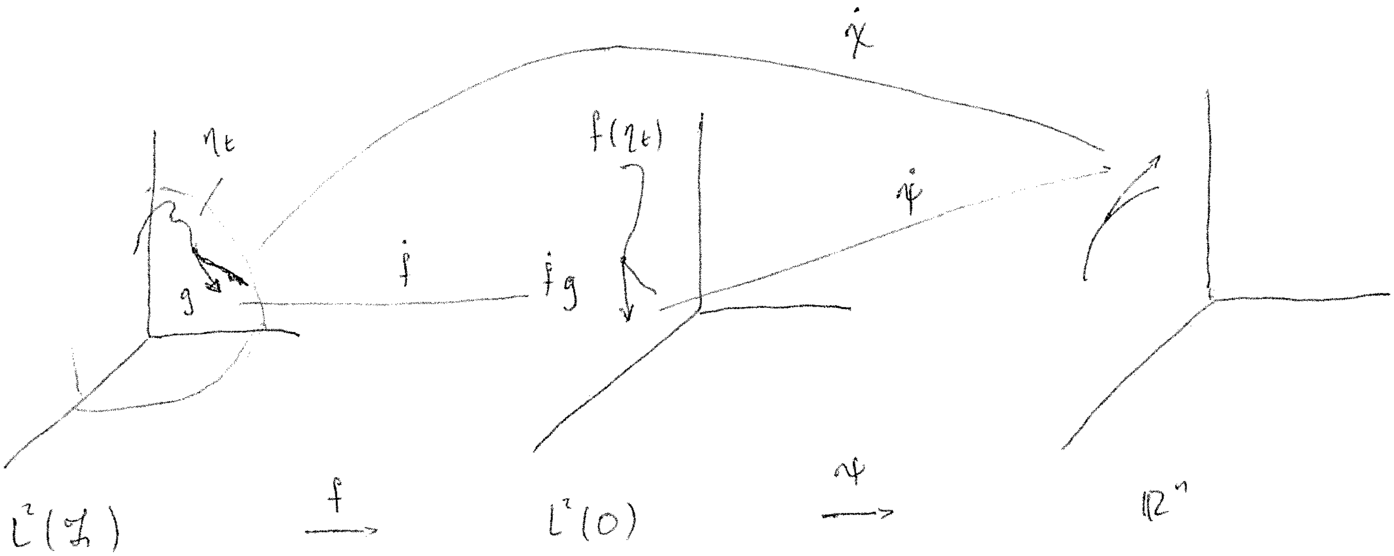
Semiparametrics

② Coarsening or "information loss model"

$$(Y^*(0), Y^*(1), X) \xrightarrow{f} (Y, A, X) = (AY^*(1) + (1-A)Y^*(0), A, X)$$

"$\mathcal{Z}$" "O"

We want influence functions of parameter on coarsened data from influence functions $\dot{\chi}$ on full data



$$\dot{\chi}(g) = (\dot{\chi}, g) = \dot{\psi}(\dot{f}g) = (\dot{\psi}, \dot{f}g) = (\dot{f}^*\dot{\psi}, g)$$

observed data influence function $\dot{\psi}$ obtained from pre-image under $\dot{f}^*$ of full-data influence function

Thm In information loss model, $\dot{f}$ is conditioning

$$\dot{f}g = \mathbb{E}(g(\mathcal{Z}) \mid O) = \mathbb{E}(g(Y^*(1), Y^*(0), X) \mid Y, A, X)$$

f^* is reverse conditioning $\dot{f}^*g = \mathbb{E}(g(O) \mid \mathcal{Z})$

ATE estimation

Strategy

1) find full-data influence functions ("χ")

2) find observed data influence functions

   a) solve $\mathbb{E}(h(Y,A,X)\mid\mathcal{Z}) = \varphi^F(\mathcal{Z})$ for h, where $\varphi^F(\mathcal{Z}) \in \Lambda_F^{\perp}$ is an unnormalized influence function

   b) solve $\mathbb{E}(h(Y,A,X)\mid\mathcal{Z}) = 0$ for h — "augmentation space"

3) find efficient influence function, derive doubly robust estimator

---

1) Non-parametric model for $\mathcal{Z} = (Y^*(0), Y^*(1), X)$, just a single influence function

$$\hat{\Psi}_n = \frac{1}{n}\sum_{j=1}^{n}\left(Y_j^*(1) - Y_j^*(0)\right), \qquad \varphi = Y^*(1) - Y^*(0) - \Psi_0$$

2a) $\mathbb{E}\left[\, h(Y,A,X) \mid Y^*(0), Y^*(1), X \,\right] = Y^*(1) - Y^*(0) - \Psi_0$

$$\Rightarrow \quad h(Y,A,X) = \frac{AY}{\pi(X)} - \frac{(1-A)Y}{1-\pi(X)} - \Psi_0$$

$$\mathbb{E}\left(\frac{AY}{\pi(X)} \,\middle|\, Y^*(0), Y^*(1), X\right) = \pi^{-1}(X)\,\mathbb{E}\left(AY^*(1) \mid \ldots\right) = Y^*(1) \ \ldots$$

2b) $\mathbb{E}\left(h(Y,A,X) \mid Y^*(0), Y^*(1), X\right) = 0$

$= \mathbb{E}\left(A\, h(Y,1,X) + (1-A)\, h(Y,0,X) \mid Y^*(0), Y^*(1), X\right)$

$= \mathbb{E}\left(A\, h(Y^*(1),1,X) + (1-A)\, h(Y^*(0),0,X) \mid Y^*(0), Y^*(1), X\right)$

$= h(Y^*(1),1,X)\,\pi(X) + h(Y^*(0),0,X)\,(1-\pi(X))$

$h(Y^*(1),1,X)\,\pi(X) = -h(Y^*(0),0,X)\,(1-\pi(X))$

$\Rightarrow h(Y,A,X) = h(A,X) = A\,h(1,X) - (1-A)\dfrac{\pi(X)}{1-\pi(X)}\,h(1,X)$

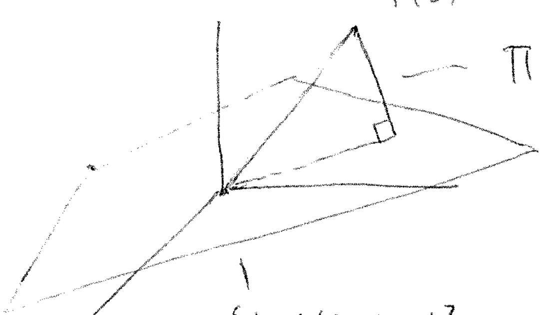
$= h(1,X)\dfrac{A-\pi(X)}{1-\pi(X)} = h'(X)\,(A-\pi(X))$ * score for $\pi_t(X)$

3) Observed data influence functions are

$$\frac{AY}{\pi(X)} - \frac{(1-A)Y}{1-\pi(X)} + h(X)(A-\pi(X)) - \Psi_0\,, \qquad h \in L^2(X)$$

minimize norm (variance) at the residual

$\dot\psi(0)$



$\Pi\left(\dot\psi \mid h(X)(A-\pi(X))\right)$

$\{h(X)(A-\pi(X))\}_h$

- linearize in A and use

$\Pi\left(A f(Y,X) \mid (A-\pi(X))h(X)\right) = (A-\pi(X))\,\mathbb{E}(f \mid A=1, X)$

- $$\psi_{eff}(Y,A,X) = \frac{AY}{\pi(X)} - \frac{(A-\pi(X))\,\mathbb{E}(Y \mid A=1,X)}{\pi(X)} - \frac{(1-A)Y}{1-\pi(X)} - \frac{(A-\pi(X))\,\mathbb{E}(Y \mid A=0,X)}{1-\pi(X)} - \Psi_0$$