Causal Inference Methods in Data Science Lecture 3: Time-Varying Treatment Effects, Marginal Structural Models, Structural Nested Models, Optimal Dynamic Treatment Regimes

Lin Liu

April 7, 2024

Some references

For the ease of presentation, we only consider two different time points. For general case, you can bootstrap from the materials of this note by induction. Also read the following papers:

Naimi, Cole, Kennedy. An introduction to g methods. International Journal of Epidemiology. 2017.

Robins. Association, causation, and marginal structural models. 1999.

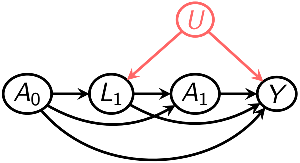
Robins. Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. 2000.

Murphy. Optimal dynamic treatment regimes. JRSS-B 2003.

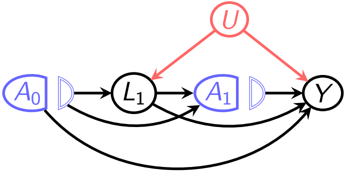
Robins. Optimal Structural Nested Models for Optimal Sequential Decisions. 2004

Zhang, Bareinboim. Designing Optimal Dynamic Treatment Regimes: A Causal Reinforcement Learning Approach. ICML 2020.

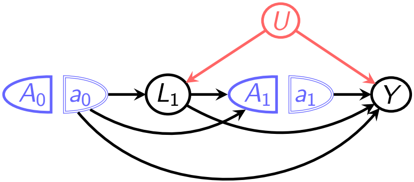
From DAG to SWIG: step by step



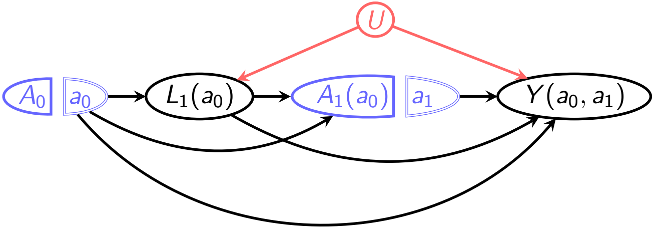
From DAG to SWIG: step by step



From DAG to SWIG: step by step

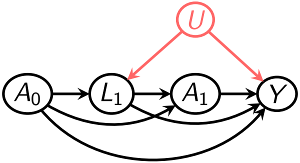


From DAG to SWIG: step by step

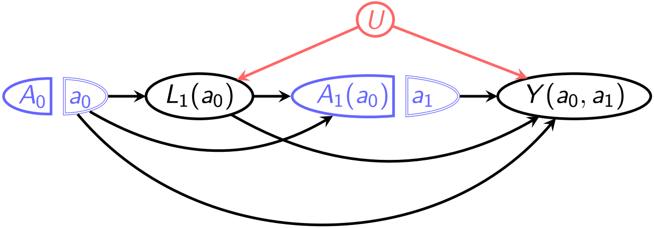


One more example of SWIG: Verma

Can we identify $\mathbb{E}[Y(a_0, a_1)]$ in the following complete Verma's graph?



SWIG?



Reading CI: $Y(a_0, a_1) \perp\!\!\!\perp A_0$ and $Y(a_0, a_1) \perp\!\!\!\perp A_1(a_0) | L_1(a_0), A_0$

Story behind Verma's graph

- Two-stage sequentially randomized trial
  - At $t = 0$, flip a coin to make a decision if $A_0 = 0$ or 1

Story behind Verma's graph

- Two-stage sequentially randomized trial
  - At $t = 0$, flip a coin to make a decision if $A_0 = 0$ or 1
  - At $t = 1$, observe a reward L_1 of the action A_0

Story behind Verma's graph

- Two-stage sequentially randomized trial
  - At $t = 0$, flip a coin to make a decision if $A_0 = 0$ or 1
  - At $t = 1$, observe a reward L_1 of the action A_0
  - Based on the historical information $\bar{H}_1 = (A_0, L_1)$ (for those familiar with stochastic processes, think of it as a filtration), make a randomized decision if $A_1 = 0$ or 1, with probability $\Pr(A_1 = 1 | A_0, L_1)$

Story behind Verma's graph

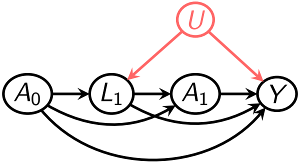
- Two-stage sequentially randomized trial
  - At $t = 0$, flip a coin to make a decision if $A_0 = 0$ or 1
  - At $t = 1$, observe a reward L_1 of the action A_0
  - Based on the historical information $\bar{H}_1 = (A_0, L_1)$ (for those familiar with stochastic processes, think of it as a filtration), make a randomized decision if $A_1 = 0$ or 1, with probability $\Pr(A_1 = 1 | A_0, L_1)$
  - Finally observe the outcome Y

Story behind Verma's graph

- Two-stage sequentially randomized trial
  - At $t = 0$, flip a coin to make a decision if $A_0 = 0$ or 1
  - At $t = 1$, observe a reward L_1 of the action A_0
  - Based on the historical information $\bar{H}_1 = (A_0, L_1)$ (for those familiar with stochastic processes, think of it as a filtration), make a randomized decision if $A_1 = 0$ or 1, with probability $\Pr(A_1 = 1 | A_0, L_1)$
  - Finally observe the outcome Y

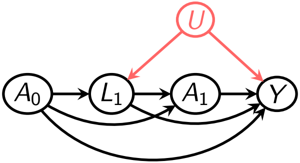
- Discuss why in a sequentially randomized trial, there could still exist an unmeasured common cause U of L_1 and Y

Time-varying treatment effects



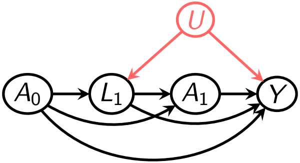
Goal: $\theta = \mathbb{E}[Y(a_0, a_1)]$ from data e.g. $\tau = \mathbb{E}[Y(1,1)] - \mathbb{E}[Y(1,0)]$

Time-varying treatment effects



Goal: $\theta = \mathbb{E}[Y(a_0, a_1)]$ from data e.g. $\tau = \mathbb{E}[Y(1,1)] - \mathbb{E}[Y(1,0)]$

Time-varying treatment effects

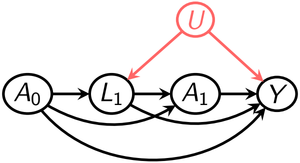


Goal: $\theta = \mathbb{E}[Y(a_0, a_1)]$ from data e.g. $\tau = \mathbb{E}[Y(1,1)] - \mathbb{E}[Y(1,0)]$

Historically, this was posed as an impenetrable problem by then famous epidemiologist E.S. Gilbert (mother of Peter Gilbert, HIV epidemiologist at UW) because:

(1) L_1 is a mediator through which the action/treatment A_0 causes the outcome Y

Time-varying treatment effects



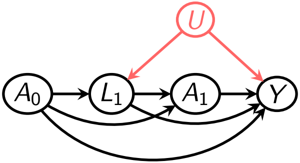
Goal: $\theta = \mathbb{E}[Y(a_0, a_1)]$ from data e.g. $\tau = \mathbb{E}[Y(1,1)] - \mathbb{E}[Y(1,0)]$

Historically, this was posed as an impenetrable problem by then famous epidemiologist E.S. Gilbert (mother of Peter Gilbert, HIV epidemiologist at UW) because:

(1) L_1 is a mediator through which the action/treatment A_0 causes the outcome Y

(2) L_1 is a confounder which causes both the action/treatment A_1 and the outcome Y

Time-varying treatment effects



Goal: $\theta = \mathbb{E}[Y(a_0, a_1)]$ from data e.g. $\tau = \mathbb{E}[Y(1,1)] - \mathbb{E}[Y(1,0)]$

Historically, this was posed as an impenetrable problem by then famous epidemiologist E.S. Gilbert (mother of Peter Gilbert, HIV epidemiologist at UW) because:

(1) L_1 is a mediator through which the action/treatment A_0 causes the outcome Y

(2) L_1 is a confounder which causes both the action/treatment A_1 and the outcome Y

(3) Feedback from Y to A_1 via U

Conventional analytical methods fail to estimate time-varying causal effect

- When E.S. Gilbert posed the problem, people only had regression in the toolbox

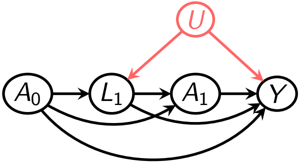
Conventional analytical methods fail to estimate time-varying causal effect

- When E.S. Gilbert posed the problem, people only had regression in the toolbox
- The question became whether one should adjust for L_1 in the regression $Y \sim A_0, A_1$

Conventional analytical methods fail to estimate time-varying causal effect

- When E.S. Gilbert posed the problem, people only had regression in the toolbox

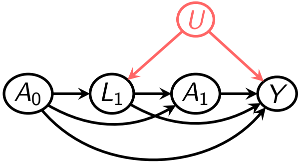
- The question became whether one should adjust for L_1 in the regression $Y \sim A_0, A_1$



- Adjusting for L_1: backdoor path $A_0 \rightarrow L_1 \leftarrow U \rightarrow Y$ is open

Conventional analytical methods fail to estimate time-varying causal effect

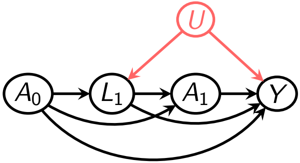
- When E.S. Gilbert posed the problem, people only had regression in the toolbox
- The question became whether one should adjust for L_1 in the regression $Y \sim A_0, A_1$



- Not adjusting for L_1: backdoor path $A_1 \leftarrow L_1 \rightarrow Y$ is open

Conventional analytical methods fail to estimate time-varying causal effect

- When E.S. Gilbert posed the problem, people only had regression in the toolbox
- The question became whether one should adjust for L_1 in the regression $Y \sim A_0, A_1$



- Regression type control fails regardless adjusting for L_1 or not

Time varying treatment effect: identification conditions

Similar to the single time point case, we have the following set of identification conditions

1. Consistency:

$$Y = \sum_{a_0, a_1} Y(a_0, a_1) \mathbb{1}\{A_0 = a_0, A_1 = a_1\}$$

$$A_1 = \sum_{a_0} A_1(a_0) \mathbb{1}\{A_0 = a_0\}$$

$$L_1 = \sum_{a_0} L_1(a_0) \mathbb{1}\{A_0 = a_0\}$$

Time varying treatment effect: identification conditions

Similar to the single time point case, we have the following set of identification conditions

1 Consistency:

$$Y = \sum_{a_0, a_1} Y(a_0, a_1) \mathbb{1}\{A_0 = a_0, A_1 = a_1\}$$

$$A_1 = \sum_{a_0} A_1(a_0) \mathbb{1}\{A_0 = a_0\}$$

$$L_1 = \sum_{a_0} L_1(a_0) \mathbb{1}\{A_0 = a_0\}$$

2 Positivity/Overlap:

$$\Pr(A_0 = a_0) > 0, \forall a_0$$

$$\Pr(A_1 = a_1 | A_0 = a_0, L_1 = \ell_1) > 0, \forall a_1, a_0, \ell_1$$

Time varying treatment effect: identification conditions

Similar to the single time point case, we have the following set of identification conditions

1 Consistency:

$$
\begin{aligned}
Y &= \sum_{a_0, a_1} Y(a_0, a_1)\mathbb{1}\{A_0 = a_0, A_1 = a_1\} \\
A_1 &= \sum_{a_0} A_1(a_0)\mathbb{1}\{A_0 = a_0\} \\
L_1 &= \sum_{a_0} L_1(a_0)\mathbb{1}\{A_0 = a_0\}
\end{aligned}
$$

2 Positivity/Overlap:

$$
\begin{aligned}
&\Pr(A_0 = a_0) > 0, \forall a_0 \\
&\Pr(A_1 = a_1 | A_0 = a_0, L_1 = \ell_1) > 0, \forall a_1, a_0, \ell_1
\end{aligned}
$$

3 Sequential ignorability/randomization:

1. $Y(a_0, a_1) \perp\!\!\!\perp A_0$
2. $Y(a_0, a_1) \perp\!\!\!\perp A_1(a_0) | L_1(a_0), A_0$

Time varying treatment effect: identification g-formula

$$
\begin{aligned}
&\mathbb{E}[Y(a_0', a_1')] \\
&\overset{3.1}{=} \mathbb{E}[Y(a_0', a_1')|A_0 = a_0'] \\
&= \mathbb{E}[\mathbb{E}[Y(a_0', a_1')|A_0 = a_0', L_1(a_0')]|A_0 = a_0'] \\
&\overset{3.2}{=} \mathbb{E}[\mathbb{E}[Y(a_0', a_1')|A_0 = a_0', L_1(a_0'), A_1(a_0') = a_1']|A_0 = a_0'] \\
&\overset{1}{=} \mathbb{E}[\mathbb{E}[Y|A_0 = a_0', L_1, A_1 = a_1']|A_0 = a_0'] \\
&= \int_y \int_{\ell_1} yf(Y = y|A_0 = a_0', L_1 = \ell_1, A_1 = a_1')f(L_1 = \ell_1|A_0 = a_0')\mathrm{d}\ell_1\mathrm{d}y
\end{aligned}
$$

From g formula to IPW: change of measure

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \mathbb{1}(a_1 = a_1') f(\ell_1|a_0) \mathbb{1}(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

Can you directly write the IPW formula now?

From g formula to IPW: change of measure

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int yf(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int yf(y|a_0, \ell_1, a_1) \mathbb{1}(a_1 = a_1') f(\ell_1|a_0) \mathbb{1}(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

Can you directly write the IPW formula now?

$$\mathbb{E}[Y(a_0', a_1')] = \mathbb{E}\left[\frac{\mathbb{1}\{A_0 = a_0'\}\mathbb{1}\{A_1 = a_1'\}}{\Pr(A_0 = a_0')\Pr(A_1 = a_1'|L_1, A_0 = a_0')} Y\right] \equiv \mathbb{E}_{ipw}[Y]$$

A more intuitive identification strategy

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \mathbb{1}(a_1 = a_1') f(\ell_1|a_0) \mathbb{1}(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- The derivation of g formula is a bit mechanical

A more intuitive identification strategy

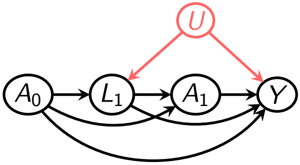
Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \mathbb{1}(a_1 = a_1') f(\ell_1|a_0) \mathbb{1}(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- The derivation of g formula is a bit mechanical
- More insightful interpretation? What change of measure does to the DAG?



A more intuitive identification strategy

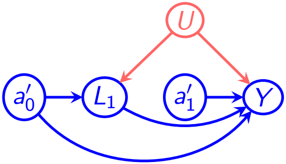
Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

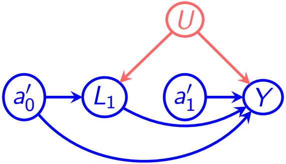
$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \mathbb{1}(a_1 = a_1') f(\ell_1|a_0) \mathbb{1}(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- The derivation of g formula is a bit mechanical
- More insightful interpretation? Cutting Arrows!



Oh, the world after change of measure!

In the IPW representation, it is obvious that we have change the distribution of observables and this new distribution (let's call it IPW distribution) is represented by



Under this new distribution, we can either use the backdoor criterion or turn this blue DAG into a blue SWIG. Either way, we can read off that no need to adjust for/control for L_1, which also explains why we just take the marginal mean of Y under the IPW distribution

$$\mathbb{E}[Y(a_0, a_1)] = \mathbb{E}_{ipw}[Y]$$

static regimes (hard intervention) vs. dynamic regimes (soft intervention)

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \delta(a_1 = a_1') f(\ell_1|a_0) \delta(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

static regimes (hard intervention) vs. dynamic regimes (soft intervention)

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \delta(a_1 = a_1') f(\ell_1|a_0) \delta(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- $A_0 = a_0', A_1 = a_1'$ is a static/fixed treatment regime/policy/strategy

static regimes (hard intervention) vs. dynamic regimes (soft intervention)

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \delta(a_1 = a_1') f(\ell_1|a_0) \delta(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- $A_0 = a_0', A_1 = a_1'$ is a static/fixed treatment regime/policy/strategy
- dynamic regime/policy/strategy? e.g. $g = (a_0', L_1)$

static regimes (hard intervention) vs. dynamic regimes (soft intervention)

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \delta(a_1 = a_1') f(\ell_1|a_0) \delta(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- $A_0 = a_0', A_1 = a_1'$ is a static/fixed treatment regime/policy/strategy
- dynamic regime/policy/strategy? e.g. $g = (a_0', L_1)$
- parameter of interest? $\mathbb{E}[Y(g)] = \mathbb{E}[Y(a_0', L_1)]$

static regimes (hard intervention) vs. dynamic regimes (soft intervention)

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \delta(a_1 = a_1') f(\ell_1|a_0) \delta(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

- $A_0 = a_0', A_1 = a_1'$ is a static/fixed treatment regime/policy/strategy
- dynamic regime/policy/strategy? e.g. $g = (a_0', L_1)$
- parameter of interest? $\mathbb{E}[Y(g)] = \mathbb{E}[Y(a_0', L_1)]$
- what is the g formula?

static regimes (hard intervention) vs. dynamic regimes (soft intervention)

Formally, g formula for $\mathbb{E}[Y(a_0', a_1')]$ is simply replacing the treatment densities in $\mathbb{E}[Y]$:

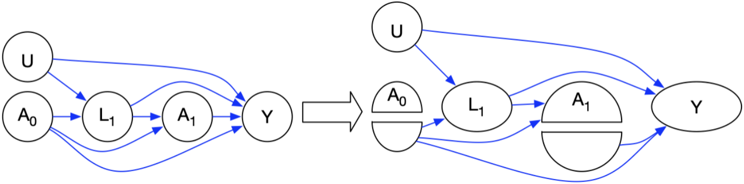
$$\mathbb{E}[Y] = \int y f(y|a_0, \ell_1, a_1) f(a_1|a_0, \ell_1) f(\ell_1|a_0) f(a_0) \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

by point mass at a_0' and a_1'

$$\mathbb{E}[Y(a_0', a_1')] = \int y f(y|a_0, \ell_1, a_1) \delta(a_1 = a_1') f(\ell_1|a_0) \delta(a_0 = a_0') \mathrm{d}a_0 \mathrm{d}\ell_1 \mathrm{d}a_1 \mathrm{d}y$$

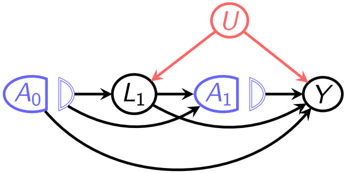
- $A_0 = a_0', A_1 = a_1'$ is a static/fixed treatment regime/policy/strategy
- dynamic regime/policy/strategy? e.g. $g = (a_0', L_1)$
- parameter of interest? $\mathbb{E}[Y(g)] = \mathbb{E}[Y(a_0', L_1)]$
- what is the g formula?
- what is the IPW formula?

Construction of SWIGs for Dynamic Regimes

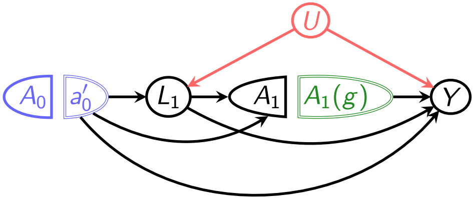


1. Split the A_0 and A_1 nodes, labeling the halves with incoming arrows A_0 and A_1 and leaving the halves with outgoing arrows unlabeled for now.
2. Check that:
   - All arrows out of the original A_0 and A_1 are now out the unlabeled halves.
   - All arrows into the original A_0 and A_1 are into the new A_0 and A_1.

From DAG to dynamic regime SWIG

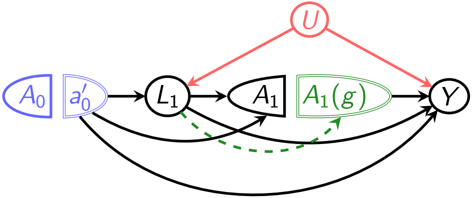


From DAG to dynamic regime SWIG



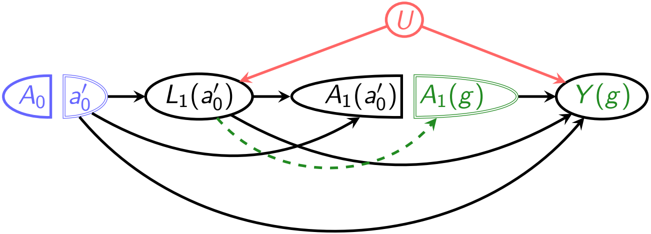
From DAG to dynamic regime SWIG

Since $A_1(g) = L_1$:



From DAG to dynamic regime SWIG

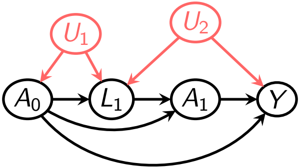
Completed!



Is $\mathbb{E}[Y(g)]$ identified? Yes! $Y(g) \perp\!\!\!\perp A_0$ and $Y(g) \perp\!\!\!\perp A_1(a_0') | L_1(a_0')$

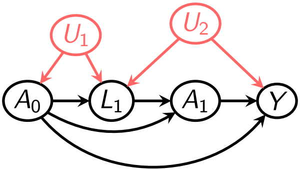
Another example

DAG

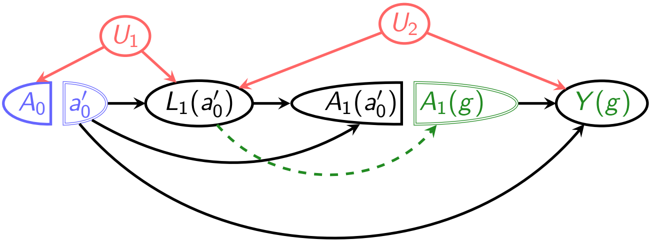


Another example

DAG



Dynamic Regime SWIG:



$Y(g) \not\perp\!\!\!\perp A_0$: can you see why? What if we change the SWIG to static regime case

Marginal Structural Models (MSM)

- We will only use one slide to explain MSM; for more details, check Hernan and Robins' textbook

Marginal Structural Models (MSM)

- We will only use one slide to explain MSM; for more details, check Hernan and Robins' textbook
- For two decision occasions, we can still estimate $\mathbb{E}[Y(a_0, a_1)]$ for each possible a_0, a_1

Marginal Structural Models (MSM)

- We will only use one slide to explain MSM; for more details, check Hernan and Robins' textbook
- For two decision occasions, we can still estimate $\mathbb{E}[Y(a_0, a_1)]$ for each possible a_0, a_1
- But what if we have $T = 100$ decision occasions?

Marginal Structural Models (MSM)

- We will only use one slide to explain MSM; for more details, check Hernan and Robins' textbook
- For two decision occasions, we can still estimate $\mathbb{E}[Y(a_0, a_1)]$ for each possible a_0, a_1
- But what if we have $T = 100$ decision occasions?
- Then we have to make complexity-reducing modeling assumptions such as

$$\mathbb{E}[Y(\bar{a})] = \boldsymbol{\beta}^\top \boldsymbol{h}(\bar{a})$$

and usually such a model is related to the scientific problem itself

connection to off-policy learning

- What is off-policy learning? Observe naturally generated data according to some distribution

$$(L_0, A_0, \cdots, L_t, A_t, \cdots, Y)$$

connection to off-policy learning

- What is off-policy learning? Observe naturally generated data according to some distribution

$$(L_0, A_0, \cdots, L_t, A_t, \cdots, Y)$$

- We are not interested in learning how $A_0, \cdots$ are assigned by nature (by the law in the observed data)

connection to off-policy learning

- What is off-policy learning? Observe naturally generated data according to some distribution

$$(L_0, A_0, \cdots, L_t, A_t, \cdots, Y)$$

- We are not interested in learning how $A_0, \cdots$ are assigned by nature (by the law in the observed data)

- We are interested in learning a new policy, e.g.

$$A_0 | L_0 = \ell_0 \sim \text{Bern}\left(\frac{\exp\{0.5\ell_0\}}{1+\exp\{0.5\ell_0\}}\right), \cdots$$

connection to off-policy learning

- What is off-policy learning? Observe naturally generated data according to some distribution

$$(L_0, A_0, \cdots, L_t, A_t, \cdots, Y)$$

- We are not interested in learning how $A_0, \cdots$ are assigned by nature (by the law in the observed data)

- We are interested in learning a new policy, e.g.

$$A_0 | L_0 = \ell_0 \sim \text{Bern}\left(\frac{\exp\{0.5\ell_0\}}{1 + \exp\{0.5\ell_0\}}\right), \cdots$$

- easy task: MSM/change of measure/density ratio/IPW

connection to off-policy learning

- What is off-policy learning? Observe naturally generated data according to some distribution

$$(L_0, A_0, \cdots, L_t, A_t, \cdots, Y)$$

- We are not interested in learning how $A_0, \cdots$ are assigned by nature (by the law in the observed data)

- We are interested in learning a new policy, e.g.

$$A_0 | L_0 = \ell_0 \sim \text{Bern}\left(\frac{\exp\{0.5\ell_0\}}{1+\exp\{0.5\ell_0\}}\right), \cdots$$

- easy task: MSM/change of measure/density ratio/IPW

- optimal dynamic treatment regimes: consider a class of possible regimes g for each action/decision occasion,

$$g^{opt} = \arg\ \max_{g \in \mathcal{G}} \mathbb{E}[Y(g)]$$

observing reward at each time point?

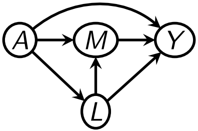
In more RL settings, we have the following data structure (Y denotes rewards)

$$(L_0, A_0, Y_0, \cdots, L_t, A_t, Y_t, \cdots, Y_T)$$

which is nothing but a longitudinal study with repeated outcome measurements (we have been only discussing longitudinal studies without repeated outcome)

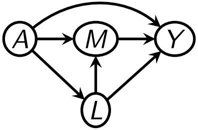
Mediation analysis revisited: treatment-induced mediator-outcome confounding

REF: Identifiability of path-specific effects. UAI 2005
REF: Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. Epidemiology 2014



Suppose we are interested in the effect of A on Y through and not through M
Recall the cross-world ID assumption: $Y(a, m) \perp\!\!\!\perp M(a')$ (since there exist no baseline confounders)

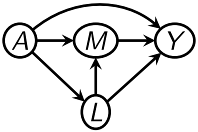
Mediation analysis revisited: treatment-induced mediator-outcome confounding



Suppose we are interested in the effect of A on Y through and not through M

Recall the cross-world ID assumption: $Y(a, m) \perp\!\!\!\perp M(a')$ (since there exist no baseline confounders)

Mediation analysis revisited: treatment-induced mediator-outcome confounding



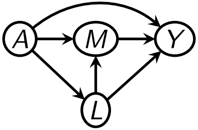
Suppose we are interested in the effect of A on Y through and not through M

Recall the cross-world ID assumption: $Y(a, m) \perp\!\!\!\perp M(a')$ (since there exist no baseline confounders)

- Due to the presence of L, $Y(a, m) \not\perp\!\!\!\perp M(a')$ but $Y(a, m) \perp\!\!\!\perp M(a')|L$: by NPSEM-IE

$$Y(a, m) = f_Y(a, L, m, \varepsilon_Y), M(a') = f_M(a', L, \varepsilon_M)$$

Mediation analysis revisited: treatment-induced mediator-outcome confounding



Suppose we are interested in the effect of A on Y through and not through M

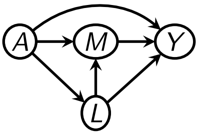
Recall the cross-world ID assumption: $Y(a, m) \perp\!\!\!\perp M(a')$ (since there exist no baseline confounders)

- Due to the presence of L, $Y(a, m) \not\perp\!\!\!\perp M(a')$ but $Y(a, m) \perp\!\!\!\perp M(a')|L$: by NPSEM-IE

$$Y(a, m) = f_Y(a, L, m, \varepsilon_Y), M(a') = f_M(a', L, \varepsilon_M)$$

- But as soon as we condition on L, the path $A \to L \to Y$ is blocked, which is a part of the effect of A on Y not through mediator M

Mediation analysis revisited: treatment-induced mediator-outcome confounding



Suppose we are interested in the effect of A on Y through and not through M

Recall the cross-world ID assumption: $Y(a,m) \perp\!\!\!\perp M(a')$ (since there exist no baseline confounders)

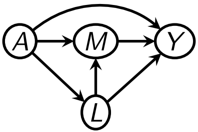
• Due to the presence of L, $Y(a,m) \not\perp\!\!\!\perp M(a')$ but $Y(a,m) \perp\!\!\!\perp M(a')|L$: by NPSEM-IE

$$Y(a,m) = f_Y(a, L, m, \varepsilon_Y), M(a') = f_M(a', L, \varepsilon_M)$$

• But as soon as we condition on L, the path $A \to L \to Y$ is blocked, which is a part of the effect of A on Y not through mediator M

• Taken together, natural direct/indirect effects (with respect to M) are unidentified

Mediation analysis revisited: treatment-induced mediator-outcome confounding



Suppose we are interested in the effect of A on Y through and not through M

Natural direct/indirect effects (with respect to M) are unidentified, but Interventional Direct/Indirect Effects can be identified even with the presence of L. Figure out why on your own.

Structural nested models

- Another class of causal models developed by Robins is Structural Nested Models (SNM)

Structural nested models

- Another class of causal models developed by Robins is Structural Nested Models (SNM)

- When $T = 1$, in a special case (Structural Nested Mean Model or SNMM), we can model the CATT (conditional average treatment effect on the treated):

$$\mathbb{E}[Y(a)|X = x, A = a] - \mathbb{E}[Y(0)|X = x, A = a] = \gamma(x, a; \psi_{true})$$

s.t. $\gamma(x, 0; \psi) = 0$ for any x, ψ and $\gamma(x, a; \psi_{true}) = 0$ if $\psi_{true} = 0$

Structural nested models

- Another class of causal models developed by Robins is Structural Nested Models (SNM)

- When $T = 1$, in a special case (Structural Nested Mean Model or SNMM), we can model the CATT (conditional average treatment effect on the treated):

$$\mathbb{E}[Y(a)|X = x, A = a] - \mathbb{E}[Y(0)|X = x, A = a] = \gamma(x, a; \psi_{true})$$

  s.t. $\gamma(x, 0; \psi) = 0$ for any x, ψ and $\gamma(x, a; \psi_{true}) = 0$ if $\psi_{true} = 0$

- If $\psi_{true} = 0$, it encodes the null hypothesis that no causal effect of A on Y

Structural nested models

- Another class of causal models developed by Robins is Structural Nested Models (SNM)

- When $T = 1$, in a special case (Structural Nested Mean Model or SNMM), we can model the CATT (conditional average treatment effect on the treated):

$$\mathbb{E}[Y(a)|X = x, A = a] - \mathbb{E}[Y(0)|X = x, A = a] = \gamma(x, a; \psi_{true})$$

s.t. $\gamma(x, 0; \psi) = 0$ for any x, ψ and $\gamma(x, a; \psi_{true}) = 0$ if $\psi_{true} = 0$

- If $\psi_{true} = 0$, it encodes the null hypothesis that no causal effect of A on Y

- A good reference:
Vansteelandt, Joffe. Structural Nested Models and G-estimation: The Partially Realized Promise. Stat. Sci. 2014.

SNMM when $T = 1$

- In SNMM, the goal is to learn about ψ_{true}, which encodes the treatment effect

SNMM when $T = 1$

- In SNMM, the goal is to learn about ψ_{true}, which encodes the treatment effect

- Create pseudo-outcome (also called mimicking counterfactuals in a paper by Judith Lok): given any ψ,

$$\widetilde{Y}_i(\psi) := Y_i - \gamma(X_i, A_i; \psi) \; \forall \; i = 1, \cdots, n$$

SNMM when $T = 1$

- In SNMM, the goal is to learn about ψ_{true}, which encodes the treatment effect

- Create pseudo-outcome (also called mimicking counterfactuals in a paper by Judith Lok): given any ψ,

$$\widetilde{Y}_i(\psi) := Y_i - \gamma(X_i, A_i; \psi) \ \forall \ i = 1, \cdots, n$$

- What counterfactual does $\widetilde{Y}(\psi_{true})$ try to mimick?

SNMM when $T = 1$

- In SNMM, the goal is to learn about ψ_{true}, which encodes the treatment effect

- Create pseudo-outcome (also called mimicking counterfactuals in a paper by Judith Lok): given any ψ,

$$\widetilde{Y}_i(\psi) := Y_i - \gamma(X_i, A_i; \psi) \ \forall \ i = 1, \cdots, n$$

- What counterfactual does $\widetilde{Y}(\psi_{true})$ try to mimick?

- By definition:

$$\begin{aligned}
\mathbb{E}[\widetilde{Y}(\psi_{true})|X, A] &= \mathbb{E}[Y|X, A] - \gamma(X, A; \psi_{true}) \\
&= \cancel{\mathbb{E}[Y(A)|X, A] - \mathbb{E}[Y(A)|X, A]} + \mathbb{E}[Y(0)|X, A] \\
&= \mathbb{E}[Y(0)|X, A]
\end{aligned}$$

SNMM when $T = 1$: What property does ψ_{true} have?

- Under no unmeasured confounding, and by tower law of expectation

$$\mathbb{E}[\widetilde{Y}(\psi_{true})|X, A] \equiv \mathbb{E}[Y(0)|X, A] \equiv \mathbb{E}[Y(0)|X] \equiv \mathbb{E}[\widetilde{Y}(\psi_{true})|X]$$

SNMM when $T = 1$: What property does ψ_{true} have?

- Under no unmeasured confounding, and by tower law of expectation

$$\mathbb{E}[\widetilde{Y}(\psi_{true})|X, A] \equiv \mathbb{E}[\widetilde{Y}(\psi_{true})|X]$$

SNMM when $T = 1$: What property does ψ_{true} have?

- Under no unmeasured confounding, and by tower law of expectation

$$\mathbb{E}[\widetilde{Y}(\psi_{true})|X, A] \equiv \mathbb{E}[\widetilde{Y}(\psi_{true})|X]$$

- Turning conditional moment constraint into marginal moment constraint: for any measurable function g,

$$\mathbb{E}\left[\left\{\widetilde{Y}(\psi_{true}) - \mathbb{E}[\widetilde{Y}(\psi_{true})|X]\right\} d(A, X)\right] \equiv 0$$

SNMM when $T = 1$: What property does ψ_{true} have?

- Under no unmeasured confounding, and by tower law of expectation

$$\mathbb{E}[\widetilde{Y}(\psi_{true})|X, A] \equiv \mathbb{E}[\widetilde{Y}(\psi_{true})|X]$$

- Turning conditional moment constraint into marginal moment constraint: for any measurable function g,

$$\mathbb{E}\left[\left\{\widetilde{Y}(\psi_{true}) - \mathbb{E}[\widetilde{Y}(\psi_{true})|X]\right\} d(A, X)\right] \equiv 0$$

- So when given i.i.d. data $\{X_i, A_i, Y_i\}_{i=1}^{n}$, we could estimate ψ_{true} by solving

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\widetilde{Y}_i(\widehat{\psi}) - \underbrace{\widehat{\mathbb{E}}[\widetilde{Y}(\widehat{\psi})|X_i]}_{\text{estimated by regression techniques}}\right\} d(A_i, X_i) = 0$$

with some user-specified choice of g with the output dimension equal to $\dim(\psi)$ (often decided by computational convenience)

Some remarks

- The moment constraint (Robins called it "G-estimation") can be made "doubly-robust":

$$\mathbb{E}\left[\left\{\widetilde{Y}(\psi_{true}) - \mathbb{E}[\widetilde{Y}(\psi_{true})|X]\right\}\{d(A,X) - \mathbb{E}[d(A,X)|X]\}\right] \equiv 0$$

- One can also use "generalized methods of moment" (GMM) from the econometrics literature to solve the above problem by solving a minimax optimization problem

$$\min_{\psi} \max_{g \in \mathcal{G}} \frac{1}{n} \sum_{i=1}^{n} \left\{\widetilde{Y}_i(\psi) - \widehat{\mathbb{E}}[\widetilde{Y}(\psi)|X_i]\right\} d(A_i, X_i)$$

- This was recently rebranded as "adversarial training/learning" by computer scientists

- Another option is to pick the g such that the estimator $\widehat{\psi}$ has the smallest variance; this strategy is often only of theoretical interest as it often leads to estimators that are hard to compute

Exercise

A very good exercise for you to gain deeper understanding of the above approach is to consider the following SNMM model:

$$X \in \{0,1\}, A \in \{0,1\}$$
$$\mathbb{E}[Y(a)|X=x, A=a] - \mathbb{E}[Y(0)|X=x, A=a] = (\psi_{true,0} + \psi_{true,1}x)a$$

Try to study the following questions:

- Is this model saturated, i.e. no more or less parameters to perfectly fit the data?

Exercise

A very good exercise for you to gain deeper understanding of the above approach is to consider the following SNMM model:

$$X \in \{0,1\}, A \in \{0,1\}$$
$$\mathbb{E}[Y(a)|X=x, A=a] - \mathbb{E}[Y(0)|X=x, A=a] = (\psi_{true,0} + \psi_{true,1}x)a$$

Try to study the following questions:

- Is this model saturated, i.e. no more or less parameters to perfectly fit the data?
- What is a reasonable choice of g? Hint: since $\dim(\psi_{true}) = 2$, $d(x,a) : \{0,1\}^2 \to \mathbb{R}^2$

Exercise

A very good exercise for you to gain deeper understanding of the above approach is to consider the following SNMM model:

$$X \in \{0,1\}, A \in \{0,1\}$$
$$\mathbb{E}[Y(a)|X=x, A=a] - \mathbb{E}[Y(0)|X=x, A=a] = (\psi_{true,0} + \psi_{true,1}x)a$$

Try to study the following questions:

- Is this model saturated, i.e. no more or less parameters to perfectly fit the data?
- What is a reasonable choice of g? Hint: since $\dim(\psi_{true}) = 2$, $d(x,a) : \{0,1\}^2 \to \mathbb{R}^2$
- With $d(x,a) = (a, ax)^\top$, derive closed-form formula of ψ_{true}

Exercise

A very good exercise for you to gain deeper understanding of the above approach is to consider the following SNMM model:

$$X \in \{0,1\}, A \in \{0,1\}$$
$$\mathbb{E}[Y(a)|X = x, A = a] - \mathbb{E}[Y(0)|X = x, A = a] = (\psi_{true,0} + \psi_{true,1}x)a$$

Try to study the following questions:

- Is this model saturated, i.e. no more or less parameters to perfectly fit the data?
- What is a reasonable choice of g? Hint: since $\dim(\psi_{true}) = 2$, $d(x,a) : \{0,1\}^2 \to \mathbb{R}^2$
- With $d(x,a) = (a, ax)^\top$, derive closed-form formula of ψ_{true}
- Combining SNMM, the formula of ψ_{true} and the assumption that both X and A are binary, derive the formula of $\mathbb{E}[Y(1)]$ and see if its formula is familiar to you

SNMM when $T = 2$

- Modeling philosophy: Blip-down the treatment effect

$$
\begin{aligned}
&\mathbb{E}[Y(a_0, a_1) - Y(a_0, 0) | X_0 = x_0, X_1 = x_1, A_0 = a_0, A_1 = a_1] \\
= &\gamma_1(x_0, x_1, a_0, a_1; \psi^{(1)}_{true}); \\
&\mathbb{E}[Y(a_0, 0) - Y(0, 0) | X_0 = x_0, A_0 = a_0] \\
= &\gamma_0(x_0, a_0; \psi^{(0)}_{true})
\end{aligned}
$$

with γ_0, γ_1 satisfying similar restrictions to $T = 1$ case

SNMM when $T = 2$

- Modeling philosophy: Blip-down the treatment effect

$$
\begin{aligned}
&\mathbb{E}[Y(a_0, a_1) - Y(a_0, 0) | X_0 = x_0, X_1 = x_1, A_0 = a_0, A_1 = a_1] \\
&= \gamma_1(x_0, x_1, a_0, a_1; \psi^{(1)}_{true}); \\
&\mathbb{E}[Y(a_0, 0) - Y(0, 0) | X_0 = x_0, A_0 = a_0] \\
&= \gamma_0(x_0, a_0; \psi^{(0)}_{true})
\end{aligned}
$$

  with γ_0, γ_1 satisfying similar restrictions to $T = 1$ case

- Under sequential randomization/ignorability/no unmeasured confounding, we construct the following mimicking counterfactuals to estimate the causal parameters ψ_{true}

$$
\begin{aligned}
\widetilde{Y}^{(1)}(\psi^{(0)}, \psi^{(1)}) &= Y - \gamma_1(X_0, X_1, A_0, A_1; \psi^{(1)}) \\
\widetilde{Y}^{(0)}(\psi^{(0)}, \psi^{(1)}) &= \widetilde{Y}^{(1)}(\psi^{(0)}, \psi^{(1)}) - \gamma_0(X_0, A_0; \psi^{(0)})
\end{aligned}
$$

opt-SNM: optimal-regime SNM, combining dynamic programming with SNM

REF: Optimal structural nested models for optimal sequential decisions. 2004 (138 pages)

opt-SNM: $T = 2$

For optimal dynamic treatment regimes, recall that our goal is to learn (say treatments are binary)

$$g_{opt} := \arg \max_{g=(g_0,g_1), g_0 : X_0 \mapsto \{0,1\}, g_1 : (X_0, A_0, X_1) \mapsto \{0,1\}} \mathbb{E}[Y(g)]$$

opt-SNM: $T = 2$

For optimal dynamic treatment regimes, recall that our goal is to learn (say treatments are binary)

$$g_{opt} := \arg \max_{g=(g_0,g_1), g_0: X_0 \mapsto \{0,1\}, g_1:(X_0,A_0,X_1)\mapsto\{0,1\}} \mathbb{E}[Y(g)]$$

opt-SNM also postulates the following model at the second occasion:

$$\mathbb{E}[Y(a_0,a_1) - Y(a_0,0) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1]$$
$$= \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)}).$$

opt-SNM: $T = 2$

For optimal dynamic treatment regimes, recall that our goal is to learn (say treatments are binary)

$$g_{opt} := \arg \max_{g=(g_0,g_1), g_0: X_0 \mapsto \{0,1\}, g_1:(X_0,A_0,X_1) \mapsto \{0,1\}} \mathbb{E}[Y(g)]$$

opt-SNM also postulates the following model at the second occasion:

$$\mathbb{E}[Y(a_0, a_1) - Y(a_0, 0) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1]$$
$$= \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)}).$$

Define

$$g_1^{opt}(x_0, a_0, x_1) := \arg \max_{a_1 \in \{0,1\}} \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)})$$

opt-SNM: $T = 2$

For optimal dynamic treatment regimes, recall that our goal is to learn (say treatments are binary)

$$g_{opt} := \arg \max_{g=(g_0,g_1), g_0: X_0 \mapsto \{0,1\}, g_1:(X_0,A_0,X_1) \mapsto \{0,1\}} \mathbb{E}[Y(g)]$$

opt-SNM also postulates the following model at the second occasion:

$$\mathbb{E}[Y(a_0,a_1) - Y(a_0,0) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1]$$
$$= \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)}).$$

Define

$$g_1^{opt}(x_0, a_0, x_1) := \arg \max_{a_1 \in \{0,1\}} \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)})$$

opt-SNM then postulates the following model at the first occasion

$$\mathbb{E}[Y(a_0, g_1^{opt}) - Y(0, g_1^{opt}) | X_0 = x_0, A_0 = a_0]$$
$$= \gamma_0^{opt}(x_0, a_0; \psi_{true}^{opt(0)})$$

opt-SNM: $T = 2$

opt-SNM also postulates the following model at the second occasion:

$$\mathbb{E}[Y(a_0, a_1) - Y(a_0, 0) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1]$$
$$= \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)}).$$

Define

$$g_1^{opt}(x_0, a_0, x_1) := \arg\max_{a_1 \in \{0,1\}} \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)})$$

opt-SNM then postulates the following model at the first occasion

$$\mathbb{E}[Y(a_0, g_1^{opt}) - Y(0, g_1^{opt}) | X_0 = x_0, A_0 = a_0]$$
$$= \gamma_0^{opt}(x_0, a_0; \psi_{true}^{opt(0)})$$

opt-SNM: $T = 2$

opt-SNM also postulates the following model at the second occasion:

$$\mathbb{E}[Y(a_0, a_1) - Y(a_0, 0) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1]$$
$$= \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)}).$$

Define

$$g_1^{opt}(x_0, a_0, x_1) := \arg \max_{a_1 \in \{0,1\}} \gamma_1^{opt}(x_0, a_0, x_1, a_1; \psi_{true}^{opt(1)})$$

opt-SNM then postulates the following model at the first occasion

$$\mathbb{E}[Y(a_0, g_1^{opt}) - Y(0, g_1^{opt}) | X_0 = x_0, A_0 = a_0]$$
$$= \gamma_0^{opt}(x_0, a_0; \psi_{true}^{opt(0)})$$

Define

$$g_0^{opt}(x_0) := \arg \max_{a_0 \in \{0,1\}} \gamma_0^{opt}(x_0, a_0; \psi_{true}^{opt(0)})$$

Try to convince yourself

the following fact: under sequential randomization and the assumption that the postulated opt-SNM is the correct model,

$$(g_0^{opt}(x_0), g_1^{opt}(x_0, a_0, x_1))^\top \equiv g^{opt}$$

where

$$g_{opt} \coloneqq \arg \max_{g=(g_0,g_1), g_0 : X_0 \mapsto \{0,1\}, g_1 : (X_0, A_0, X_1) \mapsto \{0,1\}} \mathbb{E}[Y(g)]$$

Difference between opt-SNM and Susan Murphy's work

- As I discussed in Lecture 1, Susan Murphy is the first person studying the problem of estimating optimal dynamic regimes in statistics
  REF: Optimal dynamic treatment regimes. JRSS-B Discussion Paper 2003
  In particular, the discussions by several other famous statisticians are highly recommended for students to read (e.g. Richard Gill pointed out a measure-theoretic error made by Susan Murphy)

Difference between opt-SNM and Susan Murphy's work

- As I discussed in Lecture 1, Susan Murphy is the first person studying the problem of estimating optimal dynamic regimes in statistics
  REF: Optimal dynamic treatment regimes. JRSS-B Discussion Paper 2003
  In particular, the discussions by several other famous statisticians are highly recommended for students to read (e.g. Richard Gill pointed out a measure-theoretic error made by Susan Murphy)

- Susan Murphy's framework is essentially the same as Robins' opt-SNM, with only one difference that results in different estimating strategy and the statistical properties of the estimator; see
  REF: Demystifying optimal dynamic treatment regimes. Biometrics 2007

Difference between opt-SNM and Susan Murphy's work

- For $T = 2$, Susan Murphy postulates models for the regrets

$$
\begin{aligned}
&\mathbb{E}[Y(a_0, a_1) - Y(a_0, g_1^{opt}) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1] \\
&= \gamma_1^{regret}(x_0, a_0, x_1, a_1; \psi_{true}^{regret(1)}) \\
&\mathbb{E}[Y(a_0, g_1^{opt}) - Y(g_0^{opt}, g_1^{opt}) | X_0 = x_0, A_0 = a_0] \\
&= \gamma_0^{regret}(x_0, a_0; \psi_{true}^{regret(0)})
\end{aligned}
$$

Difference between opt-SNM and Susan Murphy's work

- For $T = 2$, Susan Murphy postulates models for the regrets

$$
\begin{aligned}
&\mathbb{E}[Y(a_0, a_1) - Y(a_0, g_1^{opt})|X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1] \\
= &\gamma_1^{regret}(x_0, a_0, x_1, a_1; \psi_{true}^{regret(1)}) \\
&\mathbb{E}[Y(a_0, g_1^{opt}) - Y(g_0^{opt}, g_1^{opt})|X_0 = x_0, A_0 = a_0] \\
= &\gamma_0^{regret}(x_0, a_0; \psi_{true}^{regret(0)})
\end{aligned}
$$

- Advantage for regret modeling: many existing algorithms such as Q-learning are about regret minimization so we can stand on giant's shoulder

Difference between opt-SNM and Susan Murphy's work

- For $T = 2$, Susan Murphy postulates models for the regrets

$$
\begin{aligned}
&\mathbb{E}[Y(a_0, a_1) - Y(a_0, g_1^{opt}) | X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1] \\
= &\gamma_1^{regret}(x_0, a_0, x_1, a_1; \psi_{true}^{regret(1)}) \\
&\mathbb{E}[Y(a_0, g_1^{opt}) - Y(g_0^{opt}, g_1^{opt}) | X_0 = x_0, A_0 = a_0] \\
= &\gamma_0^{regret}(x_0, a_0; \psi_{true}^{regret(0)})
\end{aligned}
$$

- Advantage for regret modeling: many existing algorithms such as Q-learning are about regret minimization so we can stand on giant's shoulder

- Advantage for opt-SNM: at each time step, blip model only contrasts the treatment effect between a_t and "reference treatment" 0, no optimization is involved once the future optima is determined by dynamic programming

Difference between opt-SNM and Susan Murphy's work

- For $T = 2$, Susan Murphy postulates models for the regrets

$$
\begin{aligned}
&\mathbb{E}[Y(a_0, a_1) - Y(a_0, g_1^{opt})|X_0 = x_0, A_0 = a_0, X_1 = x_1, A_1 = a_1] \\
&= \gamma_1^{regret}(x_0, a_0, x_1, a_1; \psi_{true}^{regret(1)}) \\
&\mathbb{E}[Y(a_0, g_1^{opt}) - Y(g_0^{opt}, g_1^{opt})|X_0 = x_0, A_0 = a_0] \\
&= \gamma_0^{regret}(x_0, a_0; \psi_{true}^{regret(0)})
\end{aligned}
$$

- Advantage for regret modeling: many existing algorithms such as Q-learning are about regret minimization so we can stand on giant's shoulder

- Advantage for opt-SNM: at each time step, blip model only contrasts the treatment effect between a_t and "reference treatment" 0, no optimization is involved once the future optima is determined by dynamic programming

- For more detailed comparison: see page 55 of REF: Robins 2004.

O-learning framework

Finally, we introduce another framework for inferring optimal treatment regimes, which is O-learning (first developed by Michael Kosorok and colleagues)

- REF: Estimating individualized treatment rules using outcome weighted learning

O-learning framework

Finally, we introduce another framework for inferring optimal treatment regimes, which is O-learning (first developed by Michael Kosorok and colleagues)

- REF: Estimating individualized treatment rules using outcome weighted learning

- In this note, I will only describe the O-learning framework for $T = 1$ and you can generalize to $T = 2$ on your own

O-learning framework

Finally, we introduce another framework for inferring optimal treatment regimes, which is O-learning (first developed by Michael Kosorok and colleagues)

- REF: Estimating individualized treatment rules using outcome weighted learning

- In this note, I will only describe the O-learning framework for $T = 1$ and you can generalize to $T = 2$ on your own

- Key observation: for convenience, we recode the binary treatment to $\{-1, +1\}$-valued

$$g_{opt}(x) := \arg \max_{g: x \mapsto \{-1,+1\}} \mathbb{E}[Y(g)] \equiv \mathbb{E}\left[\frac{\mathbb{1}\{A = g(X)\} Y}{p(A|X)}\right]$$
$$\equiv \arg \min_{g: x \mapsto \{-1,+1\}} \underbrace{\mathbb{E}\left[\frac{Y}{p(A|X)} \mathbb{1}\{A \neq g(X)\}\right]}_{\text{weighted classification error}}$$

O-learning framework

- Of course, since indicator function is involved in the objective function, one needs to introduce surrogate losses (e.g. softmax etc.)

O-learning framework

- Of course, since indicator function is involved in the objective function, one needs to introduce surrogate losses (e.g. softmax etc.)

- But many existing theoretical results in classification can be directly applied (e.g. margin theory, VC-dimension complexity ...)

O-learning framework

- Of course, since indicator function is involved in the objective function, one needs to introduce surrogate losses (e.g. softmax etc.)
- But many existing theoretical results in classification can be directly applied (e.g. margin theory, VC-dimension complexity ...)
- A good area for optimization people to work on

Next time

Coding exercise (in R) from Lectures 2 & 3; Methods for dealing with unmeasured confounding beyond sensitivity analysis

Any Questions?