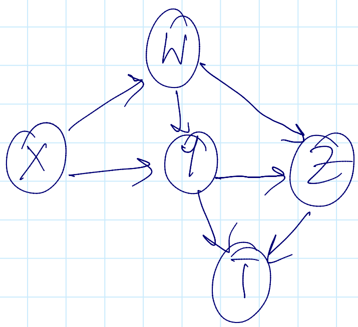
EE19MTECH01008

Sunday, 21 March 2021 10:05 PM

CS6480: Causal Inference and Learning

Assignment 1

1)



a) All parents of $Z \Rightarrow W, Y$

b) All ancestors of $Z \Rightarrow W, Y, X$

c) All children of $W \Rightarrow Y, Z$

d) All descendants of $W \Rightarrow Y, Z, T$

e) All paths between X to T →

$X \to Y \to T$, $X \to W \to Y \to T$, $X \to Y \to Z \to T$, $X \to W \to Y \to Z \to T$,

$X \to W \to Z \to T$, $X \to W \to Z \to Y \to T$, $X \to Y \to W \to Z \to T$

f) All directed paths from X to T

$X \to Y \to T$, $X \to W \to Y, T$, $X \to Y \to Z \to T$,

$X \to W \to Z \to T$, $X \to W \to Y \to Z \to T$

g) No, W & T are not marginaly independent since there are three paths through Y and Z respectively from W to T

h) $P(T, W, X, Y, Z) = P(X)\, P(W|X)\, P(Y|X)\, P(Z|W, Y, X)\, P(T|W, Y, X, Z)$

$= P(X)\, P(W|X)\, P(Y|X)\, P(Z|W, Y)\, P(T|Y, Z)$

$= P(X)\, P(W|X)\, P(Y|X)\, P(Z|W,Y)\, P(T|Y,Z)$

(by minimality & markov assumption)

i) W & T are not conditionaly independent since a path from W to T via Y still exists.

j) Yes.

$$P(W,T|Z,Y) = \frac{P(W,T,Z,Y)}{P(Z,Y)} = \sum_X \frac{P(X,W,T,Z,Y)}{P(Z,Y)}$$

$$= \sum_X \frac{P(X)\, P(W|X)\, P(Y|X,W)\, P(Z|Y,W)\, P(T|Y,Z)}{P(Z,Y)}$$

$$= \sum_X \frac{P(W,X,Y)\, P(Y,W,Z)\, P(T|Y,Z)}{P(Y,W)\, P(Z,Y)}$$

$$= P(W|Z,Y)\, P(T|Z,Y) \underbrace{\sum_X P(X|W,Y)}_{=1}$$

$$= P(W|Z,Y)\, P(T|Z,Y) \quad \text{(Proved)}$$

k) W d-separates X & T, since Y & Z are collider nodes. X & T will be independent given no conditioning on Y & Z.

2) a) $D(i,j) = 100 \times \mathbb{I}(X_{i1} \neq X_{j1}) + |X_{i2} - X_{j2}|$

$D(1,3) = 100 + |22 - 13.8| = 108.7$ $D(2,3) = |14 - 13.8| = 0.2$

$D(1,4) = |22.5 - 26.5| = 4$ $D(2,4) = |14 - 26.5| + 100 = 112.5$

$D(1,5) = |22.5 - 20| = 2.5$ $D(2,5) = 100 + |14 - 20| = 106$

$D(1,6) = |22.5 - 13.5| = 9$ $D(2,6) = 100 + |14 - 13.5| = 100.5$

$D(1,7) = |22.5 - 32.5| = 10$ $D(2,7) = 100 + |14 - 32.5| = 118.5$

$D(1,8) = 100 + |22.5 - 21| = 101.5$ $D(2,8) = |14 - 21| = 7$

We take the minimum distance for each treatment sample

We take the minimum distance for each treatement sample

∴ for 1 → best match is with 5 of the control group
since distance $D(1,5)$ is smallest among all

for 2 → best match is with 3 of the control group
since distance $D(2,3)$ is smallest among all

b) Average Treatment Effect on the treated is defined as $E[Y(1 - Y(0) \mid T=1]$

$$= E[Y(1)\ T=1] - E[Y(0) \mid T=1]$$

Here, $E[Y(0) \mid T\ 1]$ is a counterfactual variable.

Now, in randomized control trials (RCTs)
ATT is equal to ATE based on some assumptions.

i) $E[Y(0) \mid T=1] = E[Y(0) \mid T=0]$

ii) $E[Y(1) \mid T=1] = E[Y(1) \mid T=0]$

This is the exchangeability assumption.

c) $ATT = E[Y(1) - Y(0) \mid T=1]$

$$= E[Y(1) \mid T=1] - \underbrace{E[Y(0) \mid T=1]}_{\text{find using matching}}$$

∴ for 1 →,

$$E[Y(1) \mid T=1] = 30$$

$$E[Y(1) \mid T=1] = 19.5 \quad \text{(since best match is 5)}$$

$$\therefore E[Y(1) \mid T=1] - E[Y(0) \mid T=1] = 30 - 19.5$$
$$= 10.5$$

Similarly for →

$$E[Y(1) \mid T=1] = 12.5$$

$$E[Y(1) \mid T=1] = 17$$

$$E[Y(1) \mid T=1] = 17$$

$$\therefore E[Y(1) \mid T=1] - E[Y(0) \mid T=1] = 4.5$$

$$\therefore ATT = \frac{10.5 - 4.5}{2} = 3$$

5) Assuming distribution of covariates is same across treatment groups

Let W be an adjustment set

Using backdoor adjustment

$$P(y \mid do(t)) = \sum_{w} P(y \mid t, w) P(w)$$

$$= \sum_{w} \frac{P(y \mid t, w) P(t \mid w) P(w)}{P(t \mid w)}$$

$$= \sum_{w} \frac{P(y, t, w)}{P(t \mid w)}$$

Now covariates have same distribution so

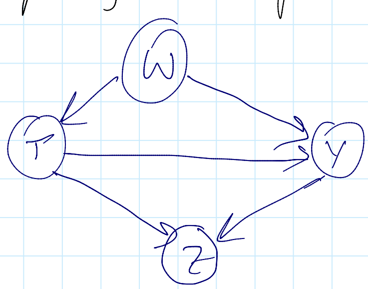
$W \perp\!\!\!\perp T$,

$$= \sum_{w} \frac{P(y, t, w)}{P(t)}$$

By Bayes rule,

$$= \sum_{w} P(y, w \mid t) = P(y \mid t)$$ (Proved)

3) Causal graph for the problem →



$E[Y \mid T=1] - E[Y \mid T=0] = 3.617$

$E[Y \mid do(T=1)] - E[Y \mid do(T=0)]$ with W adjustment $= 1.049$

$E[Y \mid do(T=1)] - E[Y \mid do(T=0)] = 0.85$ with W, Z adjustment

4) a) $Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_{1i} + \dots + u_i$

$X \to$ treatment variable
$W \to$ Control variable

By mean independence assumption we mean that $E[u|X,W] = E[u|W]$

conditional expectation of u does not depend on X.

b) $y = \beta_0 + \beta_1 x + u$

regression coefficient of x is same as causal parameter can be claimed based on mean independence assumption stated above. It makes regression an unbiased and consistent estimator.

c) $E[u] = 0$

Conditional average treatment effect (CATE)

$= E[Y^1(y) - Y^0(y))$

$= E[(\beta + \beta_1 + u) - (\beta + u)]$

$= \beta_1$

Average treatment effect $= E_{p(y)}[CATE(y)] = \beta_1$

β_1 is the coefficient of treatment variable