# Introduction to Casual Inference - 097400

# Winter 2021 - HW 3

Asaf Gendler 301727715

December 9, 2020

# Question 1:

Consider the next DAG:

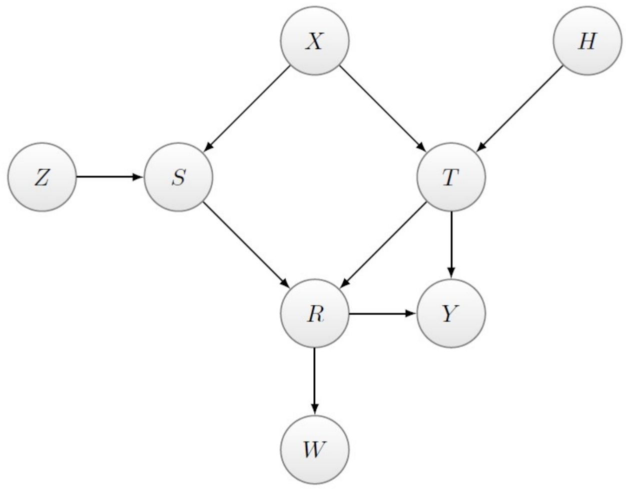


Figure 1:

## 1)

$Z \to Y$

### a)

all possible paths:

$$i : Z \to S \to R \to Y$$

$$ii : Z \to S \to R \leftarrow T \to Y$$

$$iii : Z \to S \leftarrow X \to T \to Y$$

$$iv : Z \to S \leftarrow X \to T \to R \to Y$$

### b)

to block $i$ we need to know $S$ or $R$

to block $ii$ we need to know $S$ or $T$ or to not know $R, W$ and $Y$

to block $iii$ we need to know $X$ or $T$ or to not know $S, R, W$ and $Y$

to block $iv$ we need to know $X$ or $T$ or $R$ or to not know $S, R, W$ and $Y$

| possible set of nodes to block all paths (we show only the minimal sets): $$\{S,T\},\{S,X\},\{R,T\}$$ |
|---|

## 2)

$X \to W$

### a)

all possible paths:

$$i : X \to S \to R \to W$$

$$ii : X \to T \to R \to W$$

$$iii : X \to T \to Y \leftarrow R \to W$$

### b)

to block $i$ we need to know $S$ or $R$

to block $ii$ we need to know $T$ or $R$

to block $iii$ we need to know $T$ or $R$ or to not know $Y$

| possible set of nodes to block all paths (we show only the minimal sets): $$\{R\}$$ |
|---|

## 3)

$H \to S$

### a)

all possible paths:

$$i : H \to T \leftarrow X \to S$$

$$ii : H \to T \to R \leftarrow S$$

$$iii : H \to T \to Y \leftarrow R \leftarrow S$$

**b)**

to block $i$ we need to know $X$ or to not know $T, Y, R$ and $W$

to block $ii$ we need to know $T$ or to not know $Y, R$ and $W$

to block $iii$ we need to know $T$ or $R$ or to not know $Y$

possible set of nodes to block all paths (we show only the minimal sets):

$$\emptyset$$

# Problem 2:

Consider the next casual graph:

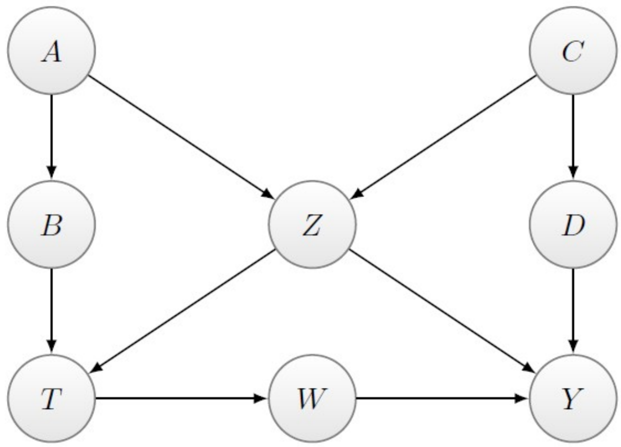


Figure 2:

## i)

all possible paths from $T$ to $Y$ that contains an arrow into $T$ are:

$$i : Y \leftarrow Z \rightarrow T$$

$$ii : Y \leftarrow Z \leftarrow A \rightarrow B \rightarrow T$$

$$iii : Y \leftarrow D \leftarrow C \rightarrow Z \rightarrow T$$

$$iv : Y \leftarrow D \leftarrow C \rightarrow Z \leftarrow A \rightarrow B \rightarrow T$$

to block $i$ we need to know $Z$

to block $ii$ we need to know $Z$ or $A$ or $B$

to block $iii$ we need to know $D$ or $C$ or $Z$

to block $iv$ we need to know $D$ or $C$ or $A$ or $B$ or to not know $Z$

in addition the sets can't contain $W$ which is a descendant of $T$

thus the sets of variables that satisfy the backdoor criteria are:

$$\{Z, A\}, \{Z, B\}, \{Z, C\}, \{Z, D\}$$

$$\{Z, A, B\}, \{Z, A, C\}, \{Z, A, D\}, \{Z, B, C\}, \{Z, B, D\}, \{Z, C, D\}$$

$$\{Z, A, B, C\}, \{Z, A, B, D\}, \{Z, A, C, D\}, \{Z, B, C, D\}$$

$$\{Z, A, B, C, D\}$$

## ii)

minimal sets of variables that satisfy the backdoor criteria:

$$\{Z, D\}, \{Z, C\}, \{Z, A\}, \{Z, B\}$$

## iii)

all possible paths from $D$ to $Y$ that contains an arrow into $D$ are:

$$i : Y \leftarrow Z \leftarrow C \rightarrow D$$

$$ii : Y \leftarrow W \leftarrow T \leftarrow Z \leftarrow C \rightarrow D$$

$$iii : Y \leftarrow W \leftarrow T \leftarrow B \leftarrow A \rightarrow Z \leftarrow C \rightarrow D$$

to block $i$ we need to know $Z$ or $C$

to block $ii$ we need to know $W$ or $T$ or $Z$ or $C$

to block $iii$ we need to know $W$ or $T$ or $B$ or $A$ or $C$ or to not know $Z$

$D$ has no descendants besides $Y$

thus the minimal set of variables that needs to be measured is:

$$\{C\}$$

# Problem 3:

## 1.

The casual graph that describes the given experiment.

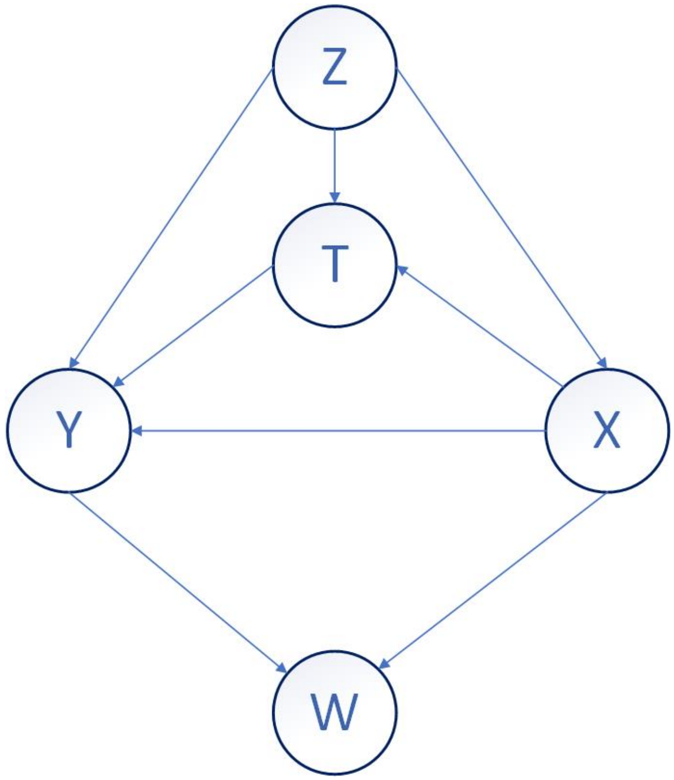


Figure 3:

## 2.

now let's estimate the conditional distributions corresponding to the graph using the data table:

| $P_Z(0)$ | $P_Z(1)$ |
|---|---|
| $\frac{9}{22}$ | $\frac{13}{22}$ |

| | $P_{X\|Z}(0)$ | $P_{X\|Z}(1)$ |
|---|---|---|
| $Z=0$ | $\frac{6}{9}=\frac{2}{3}$ | $\frac{3}{9}=\frac{1}{3}$ |
| $Z=1$ | $\frac{7}{13}$ | $\frac{6}{13}$ |

| | $P_{T\mid Z,X}(0)$ | $P_{T\mid Z,X}(1)$ |
|---|---|---|
| $Z=0,X=0$ | $\frac{2}{6}=\frac{1}{3}$ | $\frac{4}{6}=\frac{2}{3}$ |
| $Z=0,X=1$ | $\frac{1}{3}$ | $\frac{2}{3}$ |
| $Z=1,X=0$ | $\frac{3}{7}$ | $\frac{4}{7}$ |
| $Z=1,X=1$ | $\frac{4}{6}=\frac{2}{3}$ | $\frac{2}{6}=\frac{1}{3}$ |

| | $P_{W\mid X,Y}(0)$ | $P_{W\mid X,Y}(1)$ |
|---|---|---|
| $X=0,Y=0$ | $\frac{6}{8}=\frac{3}{4}$ | $\frac{2}{8}=\frac{1}{4}$ |
| $X=0,Y=1$ | $\frac{2}{5}$ | $\frac{3}{5}$ |
| $X=1,Y=0$ | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $X=1,Y=1$ | $\frac{1}{5}$ | $\frac{4}{5}$ |

| | $P_{Y\mid Z,X,T}(0)$ | $P_{Y\mid Z,X,T}(1)$ |
|---|---|---|
| $Z=0,X=0,T=0$ | 1 | 0 |
| $Z=0,X=0,T=1$ | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $Z=0,X=1,T=0$ | 1 | 0 |
| $Z=0,X=1,T=1$ | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $Z=1,X=0,T=0$ | $\frac{2}{3}$ | $\frac{1}{3}$ |
| $Z=1,X=0,T=1$ | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $Z=1,X=1,T=0$ | $\frac{1}{2}$ | $\frac{1}{2}$ |
| $Z=1,X=1,T=1$ | 0 | 1 |

## 3.

we are trying to estimate the causal effect of $T$ on $Y$

all possible paths from $T$ to $Y$ that contains an arrow into $T$ are:

$$i : T \leftarrow X \rightarrow Y$$

$$ii : T \leftarrow X \rightarrow W \leftarrow Y$$

$$iii : T \leftarrow X \leftarrow Z \rightarrow Y$$

$$iv : T \leftarrow Z \rightarrow Y$$

$$v : T \leftarrow Z \rightarrow X \rightarrow Y$$

$$vi : T \leftarrow Z \rightarrow X \rightarrow W \leftarrow Y$$

to block $i$ we need to know $X$

to block $ii$ we need to know $X$ or to not know $W$

to block $iii$ we need to know $X$ or $Z$

to block $iv$ we need to know $Z$

to block $v$ we need to know $Z$ or $X$

to block $vi$ we need to know $Z$ or $X$ or to know $W$

in addition the sets can't contain $W$ which is a descendant of $T$

thus the minimal set of variables that need to be measured in order to identify the effect of $T$ on $Y$ is:

$$\{X, Z\}$$

now using the backdoor adjustment and the distribution which we already calculated:

$$\begin{aligned} ATE &= \mathbb{E}\left[Y|do\left(T=1\right)\right] - \mathbb{E}\left[Y|do\left(T=0\right)\right] = \\ &= \mathbb{E}_{X,Z}\left[\mathbb{E}_Y\left[Y|T=1,X,Z\right]\right] - \mathbb{E}_{X,Z}\left[\mathbb{E}_Y\left[Y|T=0,X,Z\right]\right] \end{aligned}$$

$$\begin{aligned} \mathbb{E}_{X,Z}\left[\mathbb{E}_Y\left[Y|T=1,X,Z\right]\right] &= \mathbb{E}_Y\left[Y|T=1,X=0,Z=0\right]\cdot P\left(X=0,Z=0\right) + \mathbb{E}_Y\left[Y|T=1,X=0,Z=1\right]\cdot P\left(X=0,Z=1\right) \\ &+ \mathbb{E}_Y\left[Y|T=1,X=1,Z=0\right]\cdot P\left(X=1,Z=0\right) + \mathbb{E}_Y\left[Y|T=1,X=1,Z=1\right]\cdot P\left(X=1,Z=1\right) \\ &= \mathbb{E}_Y\left[Y|T=1,X=0,Z=0\right]\cdot P\left(X=0|Z=0\right)P\left(Z=0\right) + \mathbb{E}_Y\left[Y|T=1,X=0,Z=1\right]\cdot P\left(X=0|Z=1\right)P\left(Z=1\right) \\ &+ \mathbb{E}_Y\left[Y|T=1,X=1,Z=0\right]\cdot P\left(X=1|Z=0\right)P\left(Z=0\right) + \mathbb{E}_Y\left[Y|T=1,X=1,Z=1\right]\cdot P\left(X=1|Z=1\right)P\left(Z=1\right) \\ &= \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{9}{22} + \frac{1}{2}\cdot\frac{7}{13}\cdot\frac{13}{22} + \frac{1}{2}\cdot\frac{1}{3}\cdot\frac{9}{22} + 1\cdot\frac{6}{13}\cdot\frac{13}{22} \\ &= \frac{3}{22} + \frac{7}{44} + \frac{3}{44} + \frac{6}{22} = \frac{7}{11} \end{aligned}$$

$$\begin{aligned} \mathbb{E}_{X,Z}\left[\mathbb{E}_Y\left[Y|T=0,X,Z\right]\right] &= \mathbb{E}_Y\left[Y|T=0,X=0,Z=0\right]\cdot P\left(X=0,Z=0\right) + \mathbb{E}_Y\left[Y|T=0,X=0,Z=1\right]\cdot P\left(X=0,Z=1\right) \\ &+ \mathbb{E}_Y\left[Y|T=0,X=1,Z=0\right]\cdot P\left(X=1,Z=0\right) + \mathbb{E}_Y\left[Y|T=0,X=1,Z=1\right]\cdot P\left(X=1,Z=1\right) \\ &= \mathbb{E}_Y\left[Y|T=0,X=0,Z=0\right]\cdot P\left(X=0|Z=0\right)P\left(Z=0\right) + \mathbb{E}_Y\left[Y|T=0,X=0,Z=1\right]\cdot P\left(X=0|Z=1\right)P\left(Z=1\right) \\ &+ \mathbb{E}_Y\left[Y|T=0,X=1,Z=0\right]\cdot P\left(X=1|Z=0\right)P\left(Z=0\right) + \mathbb{E}_Y\left[Y|T=0,X=1,Z=1\right]\cdot P\left(X=1|Z=1\right)P\left(Z=1\right) \\ &= 0\cdot\frac{2}{3}\cdot\frac{9}{22} + \frac{1}{3}\cdot\frac{7}{13}\cdot\frac{13}{22} + 0\cdot\frac{1}{3}\cdot\frac{9}{22} + \frac{1}{2}\cdot\frac{6}{13}\cdot\frac{13}{22} \\ &= \frac{7}{66} + \frac{3}{22} = \frac{8}{33} \end{aligned}$$

$$ATE = \frac{7}{11} - \frac{8}{33} = \frac{13}{33} = 0.39$$

$$\boxed{ATE = 0.39}$$