Week 5

Wentao Gao

Causal Inference

- Probabilities and statistics

- Graphical Models and their applications

- The effect of Intervention

- Counterfactuals and Their Applications

Probabilities and statistics

Graphs

Consider the graph shown in Figure 1.8:

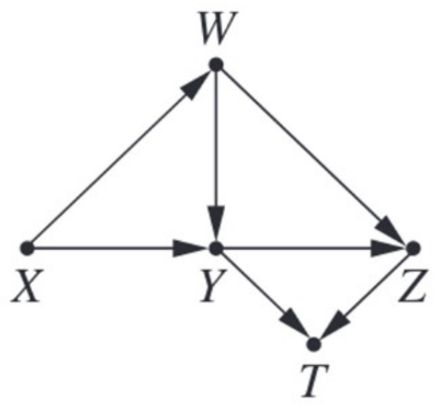


Figure 1.8 A directed graph used in Study question 1.4.1

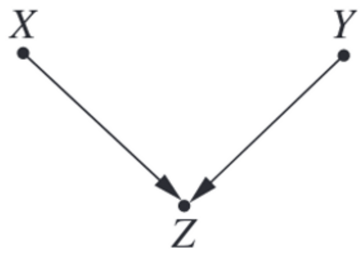
(a) Name all of the parents of Z.
(b) Name all the ancestors of Z.
(c) Name all the children of W.
(d) Name all the descendants of W.
(e) Draw all (simple) paths between X and T (i.e., no node should appear more than once).
(f) Draw all the directed paths between X and T.

Structural Causal Models

SCM 1.5.1 (Salary Based on Education and Experience)

$$U = \{X, Y\}, \quad V = \{Z\}, \quad F = \{f_Z\}$$

$$f_Z : Z = 2X + 3Y$$



SCM 2.2.5 (Temperature, Ice Cream Sales, and Crime)

$$V = \{X, Y, Z\}, U = \{U_X, U_Y, U_Z\}, F = \{f_X, f_Y, f_Z\}$$

$$f_X : X = U_X$$

$$f_Y : Y = 4x + U_Y$$

$$f_Z : Z = \frac{x}{10} + U_Z$$

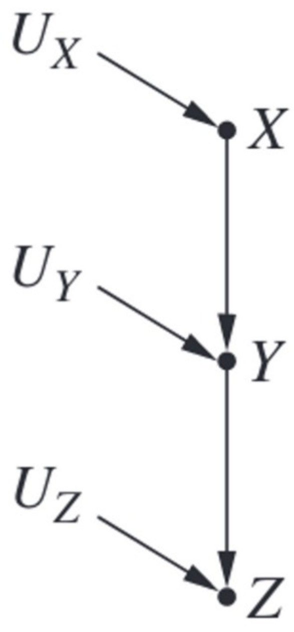
U: exogenous variables V: endogenous variables

Graphical Models and Their Applications

- Chain: A --> C --> B and A <-- C <-- B
- Fork: A <-- C --> B
- Collider: A --> C <-- B

1. Chain: A --> C --> B and A <-- C <-- B
2. Fork: A <-- C --> B
3. Collider: A --> C <-- B

Chains



SCM 2.2.1 (School Funding, SAT Scores, and College Acceptance)

$$V = \{X, Y, Z\}, U = \{U_X, U_Y, U_Z\}, F = \{f_X, f_Y, f_Z\}$$

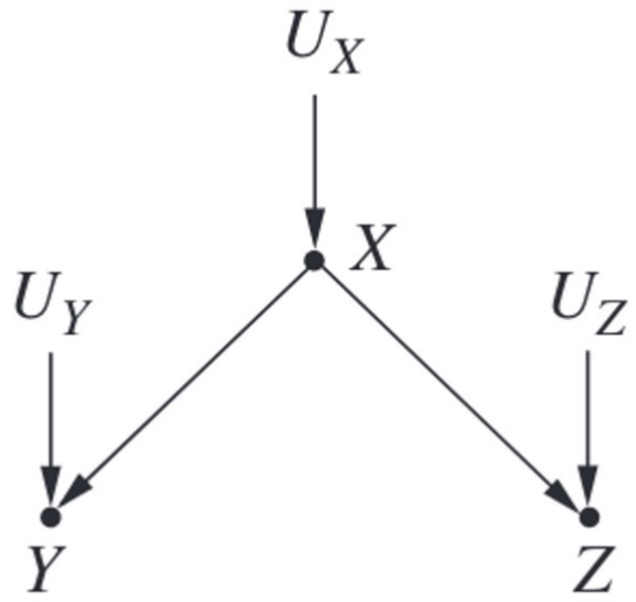
$$f_X : X = U_X$$

$$f_Y : Y = \frac{x}{3} + U_Y$$

$$f_Z : Z = \frac{y}{16} + U_Z$$

1. ***Z and Y are likely dependent***
   For some z, y, $P(Z = z|Y = y) \neq P(Z = z)$
2. ***Y and X are likely dependent***
   For some y, x, $P(Y = y|X = x) \neq P(Y = y)$
3. ***Z and X are likely dependent***
   For some z, x, $P(Z = z|X = x) \neq P(Z = z)$
4. ***Z and X are independent, conditional on Y***
   For all x, y, z, $P(Z = z|X = x, Y = y) = P(Z = z|Y = y)$

Forks



SCM 2.2.5 (Temperature, Ice Cream Sales, and Crime)

$$V = \{X, Y, Z\}, U = \{U_X, U_Y, U_Z\}, F = \{f_X, f_Y, f_Z\}$$

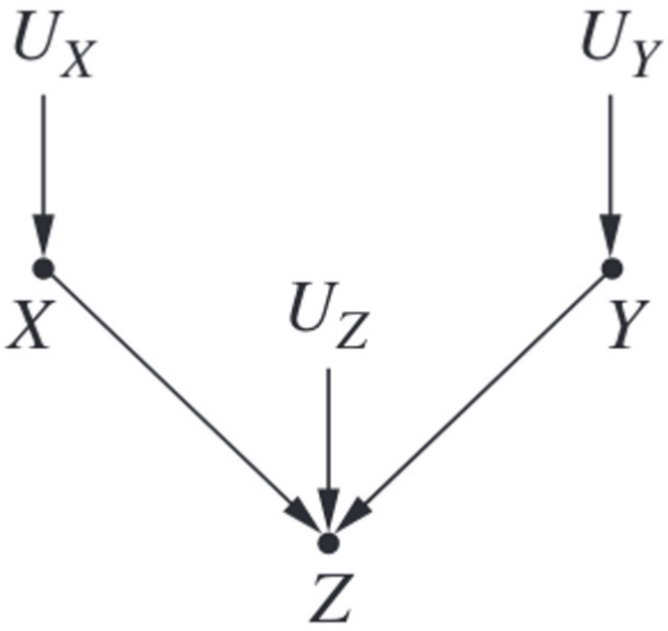
$$f_X : X = U_X$$

$$f_Y : Y = 4x + U_Y$$

$$f_Z : Z = \frac{x}{10} + U_Z$$

1. ***X and Y are likely dependent.***
   For some x, y, $P(X = x|Y = y) \neq P(X = x)$
2. ***X and Z are likely dependent.***
   For some x, z, $P(X = x|Z = z) \neq P(X = x)$
3. ***Z and Y are likely dependent.***
   For some z, y, $P(Z = z|Y = y) \neq P(Z = z)$
4. ***Y and Z are independent, conditional on X.***
   For all x, y, z, $P(Y = y|Z = z, X = x) = P(Y = y|X = x)$

Colliders



1. ***X and Z are likely dependent.***
   For some x, z, $P(X = x|Z = z) \neq P(X = x)$
2. ***Y and Z are likely dependent.***
   For some y, z, $P(Y = y|Z = z) \neq P(Y = y)$
3. ***X and Y are independent.***
   For all x, y, $P(X = x|Y = y) = P(X = x)$
4. ***X and Y are likely dependent conditional on Z.***
   For some x, y, z, $P(X = x|Y = y, Z = z) \neq P(X = x|Z = z)$

D-Separation

Definition 2.4.1 **(*d*-separation)** *A path p is blocked by a set of nodes Z if and only if*

1. *p contains a chain of nodes* $A \to B \to C$ *or a fork* $A \leftarrow B \to C$ *such that the middle node B is in Z (i.e., B is conditioned on), or*
2. *p contains a collider* $A \to B \leftarrow C$ *such that the collision node B is not in Z, and no descendant of B is in Z.*

If Z blocks every path between two nodes X and Y, then X and Y are d-separated, conditional on Z, and thus are independent conditional on Z.

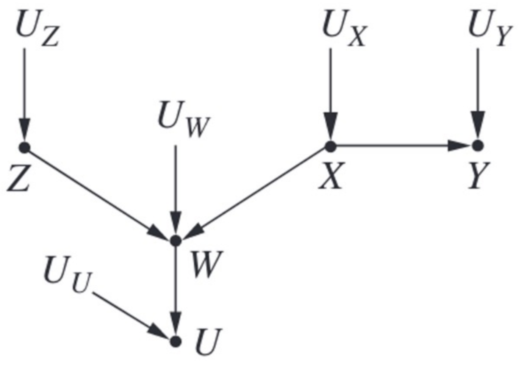


Figure 2.7 A graphical model containing a collider with child and a fork

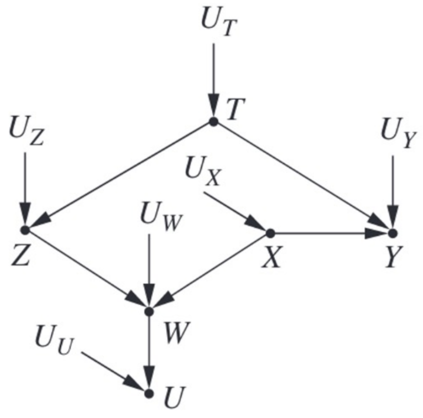


Figure 2.8 The model from Figure 2.7 with an additional forked path between *Z* and *Y*

The Effects of Interventions

Intervention

If we intervening on a variable X, then every arrow to X will be deleted.

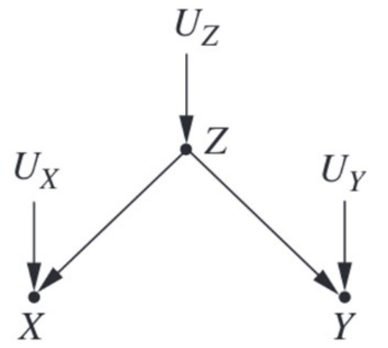


Figure 3.1 A graphical model representing the relationship between temperature (Z), ice cream sales (X), and crime rates (Y)

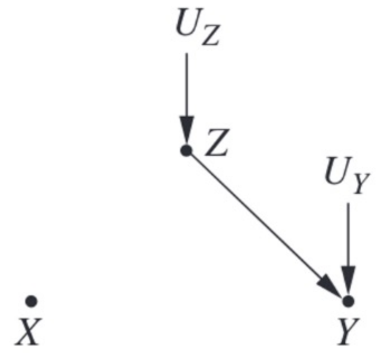


Figure 3.2 A graphical model representing an intervention on the model in Figure 3.1 that lowers ice cream sales

The Adjustment Formula

$$P(Y = 1|do(X = 1)) - P(Y = 1|do(X = 0)) \tag{3.1}$$

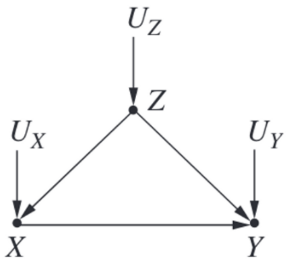


Figure 3.3 A graphical model representing the effects of a new drug, with Z representing gender, X standing for drug usage, and Y standing for recovery

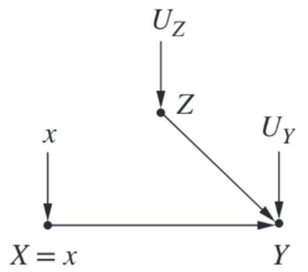


Figure 3.4 A modified graphical model representing an intervention on the model in Figure 3.3 that sets drug usage in the population, and results in the manipulated probability P_m

$$P_m(Y = y|Z = z, X = x) = P(Y = y|Z = z, X = x) \quad \text{and} \quad P_m(Z = z) = P(Z = z)$$

We can also use the fact that Z and X are d-separated in the modified model and are, therefore, independent under the intervention distribution. This tells us that Pm(Z = z|X = x)= Pm(Z = z)=P(Z = z), the last equality following from above. Putting these considerations together, we have

$$P(Y = y|do(X = x))$$

$$= P_m(Y = y|X = x) \qquad \text{(by definition)} \tag{3.2}$$

$$= \sum_z P_m(Y = y|X = x, Z = z)P_m(Z = z|X = x) \tag{3.3}$$

$$= \sum_z P_m(Y = y|X = x, Z = z)P_m(Z = z) \tag{3.4}$$

Finally, using the invariance relations, we obtain a formula for the causal effect, in terms of preintervention probabilities:

$$P(Y = y|do(X = x)) = \sum_z P(Y = y|X = x, Z = z)P(Z = z) \tag{3.5}$$

The Backdoor Criterion

Definition 3.3.1 (The Backdoor Criterion) *Given an ordered pair of variables (X, Y) in a directed acyclic graph G, a set of variables Z satisfies the* backdoor criterion *relative to (X, Y) if no node in Z is a descendant of X, and Z blocks every path between X and Y that contains an arrow into X.*

If a set of variables Z satisfies the backdoor criterion for X and Y, then the causal effect of X on Y is given by the formula

$$P(Y = y|do(X = x)) = \sum_{z} P(Y = y|X = x, Z = z)P(Z = z)$$

just as when we adjust for $PA(X)$. (Note that $PA(X)$ always satisfies the backdoor criterion.)

The logic behind the backdoor criterion is fairly straightforward. In general, we would like to condition on a set of nodes Z such that

1. We block all spurious paths between X and Y.
2. We leave all directed paths from X to Y unperturbed.
3. We create no new spurious paths.

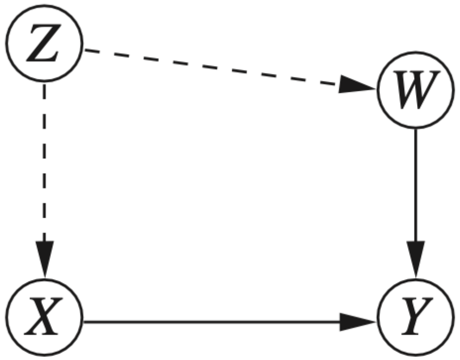


Figure 3.6 A graphical model representing the relationship between a new drug (X), recovery (Y), weight (W), and an unmeasured variable Z (socioeconomic status)

Here we are trying to gauge the effect of a drug (X) on recovery (Y). We have also measured weight (W), which has an effect on recovery. Further, we know that socioeconomic status (Z) affects both weight and the choice to receive treatment—but the study we are consulting did not record socioeconomic status.

Instead, we search for an observed variable that fits the backdoor criterion from X to Y. A brief examination of the graph shows that W, which is not a descendant of X, also blocks the backdoor path $X \leftarrow Z \rightarrow W \rightarrow Y$. Therefore, W meets the backdoor criterion. So long as the causal story conforms to the graph in Figure 3.6, adjusting for W will give us the causal effect of X on Y. Using the adjustment formula, we find

$$P(Y = y|do(X = x)) = \sum_{w} P(Y = y|X = x, W = w)P(W = w)$$

This sum can be estimated from our observational data, so long as W is observed.

The Front-Door Criterion

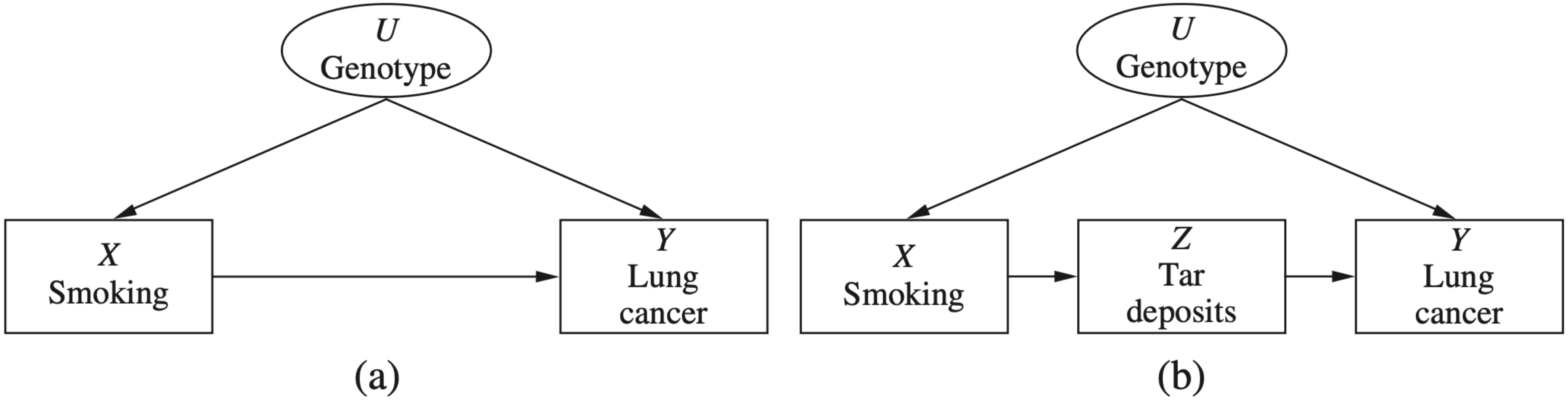


Figure 3.10 A graphical model representing the relationships between smoking (X) and lung cancer (Y), with unobserved confounder (U) and a mediating variable Z

The Front-Door Criterion

Definition 3.4.1 (Front-Door) *A set of variables Z is said to satisfy the front-door criterion relative to an ordered pair of variables (X, Y) if*

1. *Z intercepts all directed paths from X to Y.*
2. *There is no backdoor path from X to Z.*
3. *All backdoor paths from Z to Y are blocked by X.*

Theorem 3.4.1 (Front-Door Adjustment) *If Z satisfies the front-door criterion relative to (X, Y) and if $P(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by the formula*

$$P(y|do(x)) = \sum_z P(z|x) \sum_{x'} P(y|x', z)P(x') \tag{3.16}$$

The conditions stated in Definition 3.4.1 are overly conservative; some of the paths excluded by conditions (2) and (3) can actually be allowed provided they are blocked by some variables. There is a powerful symbolic machinery, called the *do-calculus*, that allows analysis of such intricate structures. In fact, the *do*-calculus uncovers *all* causal effects that can be identified from a given graph. Unfortunately, it is beyond the scope of this book (see Tian and Pearl 2002, Shpitser and Pearl 2008, Pearl 2009, and Bareinboim and Pearl 2012 for details). But the combination of the adjustment formula, the backdoor criterion, and the front-door criterion covers numerous scenarios. It proves the enormous, even revelatory, power that causal graphs have in not merely representing, but actually discovering causal information.

Inverse Probability Weighing

$$P(Y = y, Z = z | X = x) = \frac{P(Y = y, Z = z, X = x)}{P(X = x)}$$

$$P(y | do(x)) = \sum_{z} P(Y = y | X = x, Z = z) P(Z = z)$$

$$P(y | do(x)) = \sum_{z} \frac{P(Y = y | X = x, Z = z) P(X = x | Z = z) P(Z = z)}{P(X = x | Z = z)}$$

$$P(y | do(x)) = \sum_{z} \frac{P(Y = y, X = x, Z = z)}{P(X = x | Z = z)}$$

Counterfactuals and Their Applications

If we try to express this estimate using *do*-expressions, we come to an impasse. Writing

$$E(driving\ time|do(freeway), driving\ time = 1\ hour)$$

1. Actual driving time
2. Hypothetical driving time under freeway conditions when actual surface driving time is known to be 1 hour.

We denote the freeway driving time by *Y_X*=1 (or *Y*_1, where context permits) and Sepulveda driving time by *Y_X*=0 (or *Y*_0). In our case, since *Y*_0 is the *Y* actually observed, the quantity we wish to estimate is

$$E(Y_{X=1}|X = 0, Y = Y_0 = 1)$$

Defining and Computing Counterfactuals

A simple model:

$$X = aU$$

$$Y = bX + U$$

Modified model M_x:

$$X = x$$

$$Y = bX + U$$

$$U = u$$

$$Y_x(u) = bx + u$$

We begin with a fully specified model M, for which we know both the functions $\{F\}$ and the values of all exogenous variables U .
In such a deterministic model, every assignment $U = u$ to the exogenous variables corresponds to a single member of, or "unit" in a population, or to a "situation" in nature.

$Y_x(u)$ stands for *"Y would be y had X been x, in situation U = u"*

The Fundamental Law of Counterfactuals

Let M_x stand for the modified version of M, with the equation of X replaced by $X = x$. The formal definition of the counterfactual $Y_x(u)$ reads

$$Y_x(u) = Y_{M_x}(u) \tag{4.5}$$

If X is binary, then the consistency rule takes the convenient form:

$$Y = XY_1 + (1 - X)Y_0$$

From Population Data to Individual Behavior—An Illustration

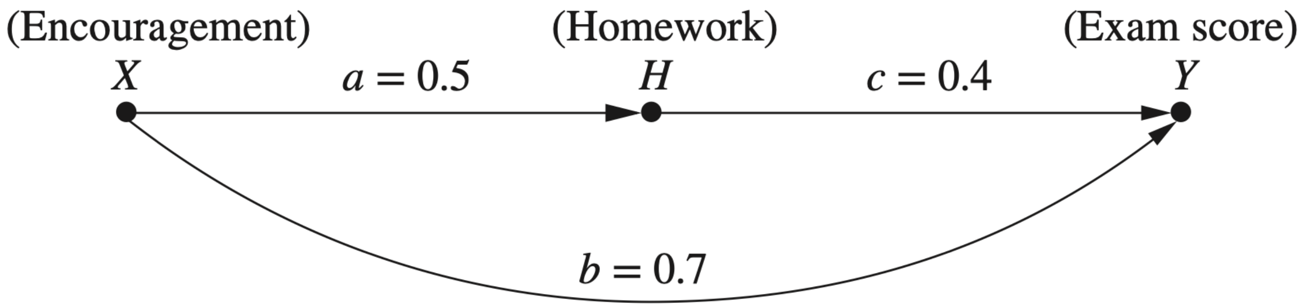


Figure 4.1 A model depicting the effect of Encouragement (X) on student's score

$$X = U_X$$
$$H = a \cdot X + U_H$$
$$Y = b \cdot X + c \cdot H + U_Y$$
$$\sigma_{U_i U_j} = 0 \quad \text{for all } i, j \in \{X, H, Y\}$$

We assume that all U factors are independent and that we are given the values for the coefficients of the Model (these can be estimated from population data):

$$a = 0.5, \quad b = 0.7, \quad c = 0.4$$

$$U_X = 0.5,$$
$$U_H = 1 - 0.5 \cdot 0.5 = 0.75, \text{ and}$$
$$U_Y = 1.5 - 0.7 \cdot 0.5 - 0.4 \cdot 1 = 0.75.$$

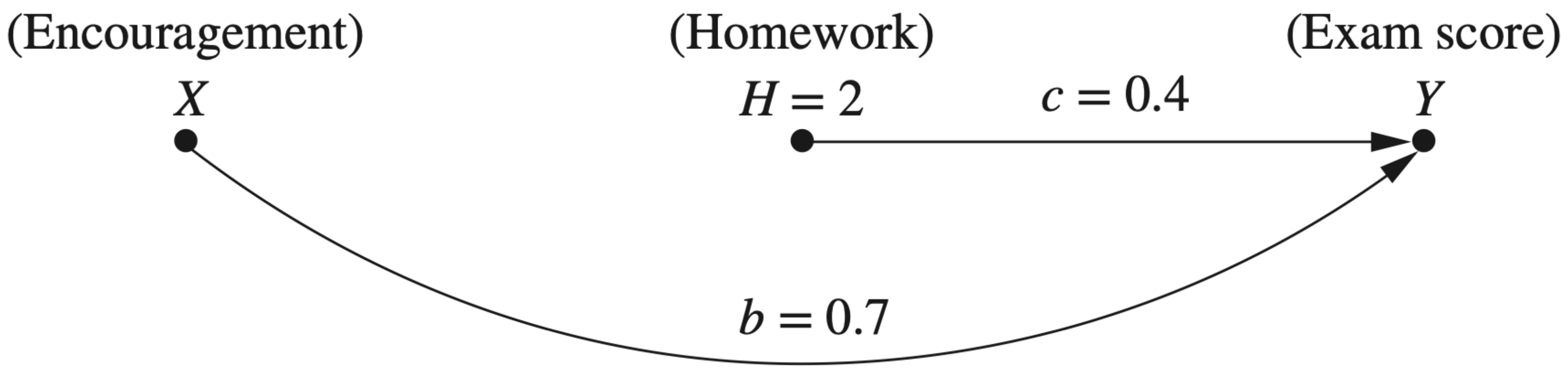


Figure 4.2 Answering a counterfactual question about a specific student's score, predicated on the assumption that homework would have increased to $H = 2$

$$
\begin{aligned}
Y_{H=2}(U_X = 0.5, U_H = 0.75, U_Y = 0.75) \\
&= 0.5 \cdot 0.7 + 2.0 \cdot 0.4 + 0.75 \\
&= 1.90
\end{aligned}
$$

The Three Steps in Computing Counterfactuals

Deterministic counterfactual:

(i) Abduction: Use evidence $E = e$ to determine the value of U.
(ii) Action: Modify the model, M, by removing the structural equations for the variables in X and replacing them with the appropriate functions $X = x$, to obtain the modified model, M_x.
(iii) Prediction: Use the modified model, M_x, and the value of U to compute the value of Y, the consequence of the counterfactual.

Given an arbitrary counterfactual of the form, $E[Y_{X=x}|E=e]$, the three-step process reads:

Nondeterministic Counterfactuals

(i) **Abduction**: Update $P(U)$ by the evidence to obtain $P(U|E = e)$.
(ii) **Action**: Modify the model, M, by removing the structural equations for the variables in X and replacing them with the appropriate functions $X = x$, to obtain the modified model, M_x.
(iii) **Prediction**: Use the modified model, M_x, and the updated probabilities over the U variables, $P(U|E = e)$, to compute the expectation of Y, the consequence of the counterfactual.

Some thinking

- The book of why introduce the level of AI Judea think. And I was convinced by the idea and thoughts of Judea. I think Causal inference is the true AI and it will be the real path of AI go.

  However, it seems that there are two kinds of AI have a better progress for now, which is Open AI using large model Chatgpt and DeepMind using RL.

- They all using the data itself but not figure out the core of the motion. I believe there will be a way to "Causality is all you need".