统计中的因果推断

沈华伟
中国科学院计算技术研究所

为什么学习因果推断

- **克服传统统计方法的不足**
  - **相关不意味着因果**（association does not imply causation）
    - 例如：用电量和冰糕销量正相关，但没有因果关系
  - **缺少从数据中解读因果关系的有效数学语言或工具**
    - 同样的数据，可以解读出不同的、甚至矛盾的结论
      - 例如：吸烟提高肺癌的患病率吗?
  - **统计无法回答反事实问题**
    - 例如：一位硕士生毕业十年后说，如果当年我转博了，现在会怎样?

J. Pearl, M. Glymour, N. P. Jewell. Causal Inference in Statistics: A Primer. Wiley, 2016.

辛普森悖论（Simpson’s Paradox）

- **总体数据上得出的统计结论和分组数据上的统计结论相反**

- **未分组**
  - 吸烟者的成绩比不吸烟者的成绩好
- **按照年龄分组后**
  - 每个分组中，吸烟者的成绩**都**比不吸烟者差
- **继续细分，在年龄的基础上再考虑收入**
  - 结论有可能会进一步反转

辛普森悖论的例子

- **关于某种药物治疗效果的数据**
  - **治疗效果：服药后是否康复**

	服药	不服药
男	87人中81人康复（**93%**）	270人234人康复（**87%**）
女	263人中192人康复（**73%**）	80人中55人康复（**69%**）
合计	350人中273人康复（**78%**）	350人中289人康复（**83%**）

✓ **总体数据来看，服药的康复率低于不服药的康复率**

✓ **按照性别分组，每个分组上服药的康复率均高于不服药的康复率**

问题出在了哪里？

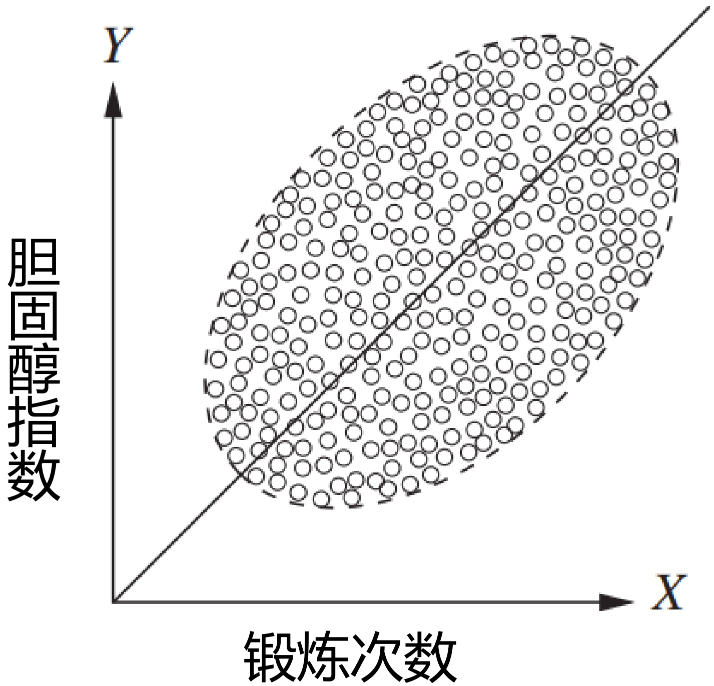
- **我们只观测到了数据，对数据背后的产生机制不清楚**

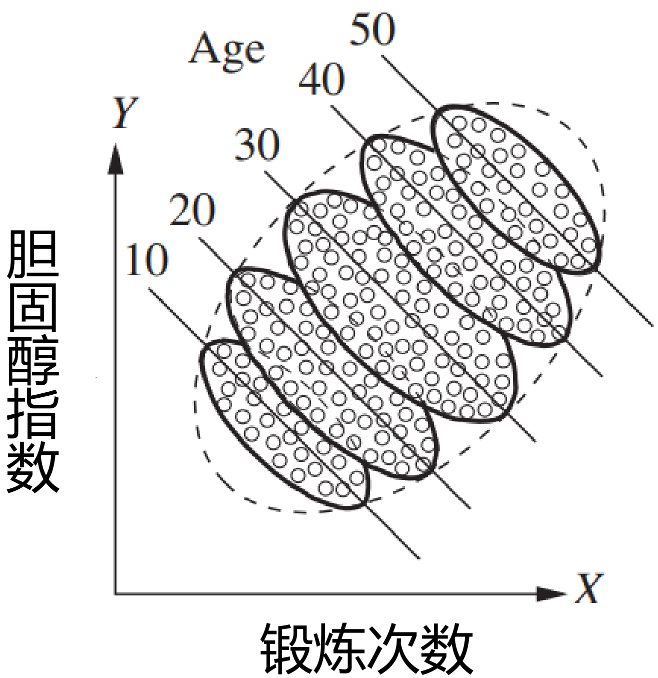
- **假如我们了解如下机制**
  - 雌性激素对于康复有副作用：女性比男性不易康复
- **解读**
  - **总体上，服药的康复率低于不服药的康复率**
    - 女性更愿意服药，样本中女性比例更高（相对于随机）
  - **按性别分组后，服药的康复率高于不服药的康复率**
    - 排除了雌性激素的影响

问题所在：性别是导致服药和康复的共同原因

再看一个类似的例子

- 锻炼次数和胆固醇指数的关系





年龄高是锻炼次数高和胆固醇指数高的的共同原因

» 再看一个不同的例子

- 和第一个例子中的数据完全相同，只是将分组依据由性别换成了血压。

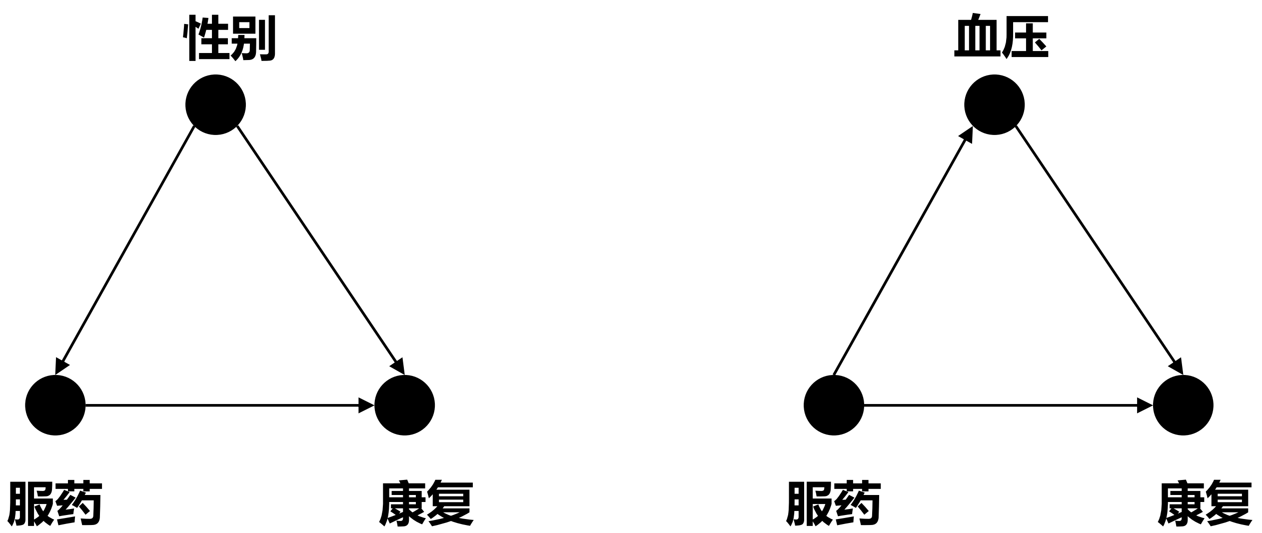
	不服药	服药
低血压	87人中81人康复（**93%**）	270人234人康复（**87%**）
高血压	263人中192人康复（**73%**）	80人中55人康复（**69%**）
合计	350人中273人康复（**78%**）	350人中289人康复（**83%**）

✓ **总体数据来看，服药的康复率高于不服药的康复率**

✓ **按照血压分组，每个分组上服药的康复率均低于不服药的康复率**

同样的数据为什么结论不同呢?

- 不同的结论：
  - 例子一：应该分组
  - 例子二：不应该分组
- 原因：**数据背后的产生机制不同**



目录

- **基础知识**
  - **结构因果模型**
  - **因果模型图**
- **干预**
  - **干预与校正公式**
  - **后门准则**
  - **前门准则**
- **反事实**
  - **反事实定义**
  - **确定性反事实计算**
  - **非确定性反事实计算**

结构因果模型 (SCM: Structural Causal Model)

- 结构因果模型用于描述数据的产生机制

- 结构因果模型的构成
  - 外生变量集合：U
    - 外生变量不依赖于其他变量
  - 内生变量集合：V
    - 内生变量至少依赖一个变量
  - 确定内生变量取值的函数集合：F

例子：教育、工作经验与薪酬水平

- **结构因果模型**

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

$$f_Z\colon Z=2X+3Y$$

X：受教育年数；Y：工作年数；Z：薪酬水平

f_Z刻画了薪酬水平和受教育年数、工作年数的关系

因果模型图

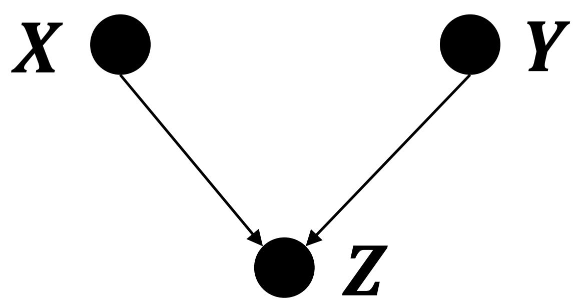
因果模型图以图的方式更直观地表示结构因果模型

结构因果模型

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

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

因果模型图



- ✓ **因果模型图中，节点表示变量，边表示变量间的依赖关系**
- ✓ **因果模型图是一个有向无环图（DAG：Directed Acyclic Graphs）**

因果模型图

- 每个结构因果模型对应一个因果模型图

- 因果模型图刻画了变量间的关系，但没有给出依赖关系的具体形式（f_Z）

- 因果模型图的优点
  - 提供了对因果关系更直观的理解！
  - 能用来非常有效地表示联合分布！

因果模型图

- 通过因果模型图及**乘积分解法则**，可以有效地表示变量联合分布

乘积分解法则： $\boldsymbol{P(x_1, x_2, \ldots, x_n) = \Pi_i P(x_i | pa_i)}$

- pa_i 表示变量 x_i 的所有父节点

- 例子：

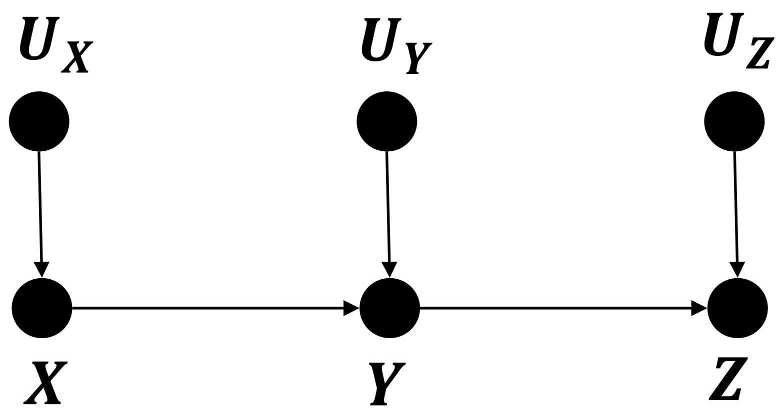
$$X \rightarrow Y \rightarrow Z$$

$$P(X=x, Y=y, Z=z) = P(X=x)P(Y=y|X=x)P(Z=z|Y=y)$$

✓ 可将"高维"分布估计问题转为一些“低维”分布估计问题，避免“维数灾难”

因果模型图与独立性

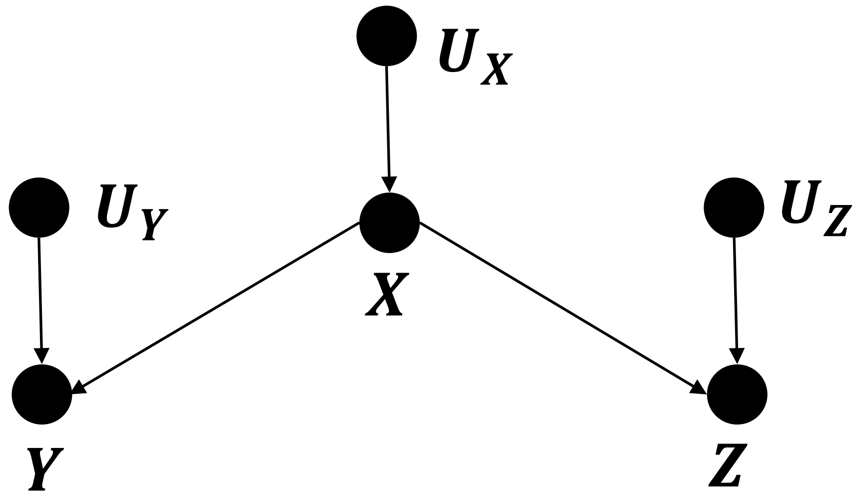
- **链结构**



  - **例子：X——开关，Y——电路，Z——灯泡状态**
- **链结构中的条件独立性： $X \perp Z \mid Y$**
  - $X = F_X(U_X)$， $Z = F_Z(Y, U_Z)$
  - 给定 Y 时， Z 只受 U_Z 影响， X 只受 U_X 影响，故X、Z 独立（已知外生变量 U_x和U_z独立）

因果模型图与独立性

- **分叉结构**

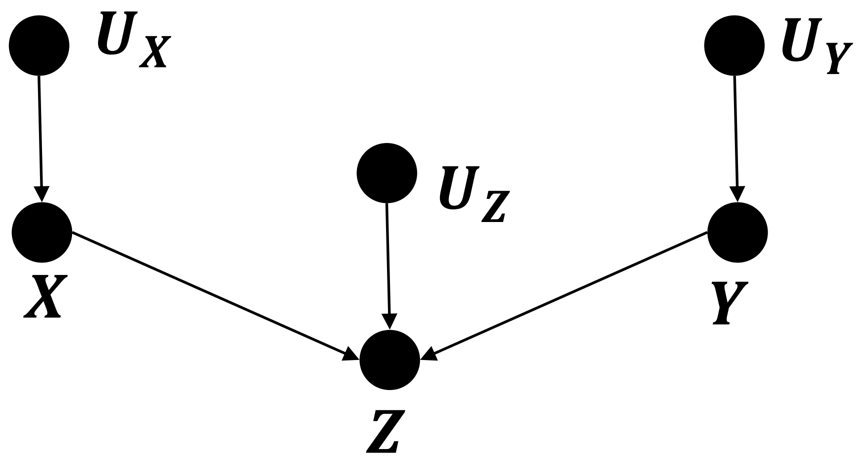


  - **例子：X——温度， Y——冰淇淋销量， Z——犯罪率**

- **分叉结构中的条件独立性： $Y \perp Z \mid X$**
  - $Y = F_Y(X, U_Y)$， $Z = F_Z(X, U_Z)$
  - 给定 X 时， Y 只受 U_Y 影响，Z 只受 U_Z 影响，故Y、Z独立（已知外生变量 U_y和U_z独立）

因果模型图与独立性

- **对撞结构**



  - **例子：X——音乐天赋， Y——学业成绩， Z——奖学金**

- **对撞结构中的条件独立性： $X \perp Y$**
  - **但在给定 Z 或Z的子孙时， X 与 Y 不独立（相互依赖）**
    - 某大学为两类学生提供奖学金（Z）：具有音乐天赋（X）或学业成绩优秀（Y）
    - 已知某学生获得了奖学金（$Z=1$），如果这个人缺乏音乐天赋（$X=0$）那么必然学业成绩优秀（$Y=1$）

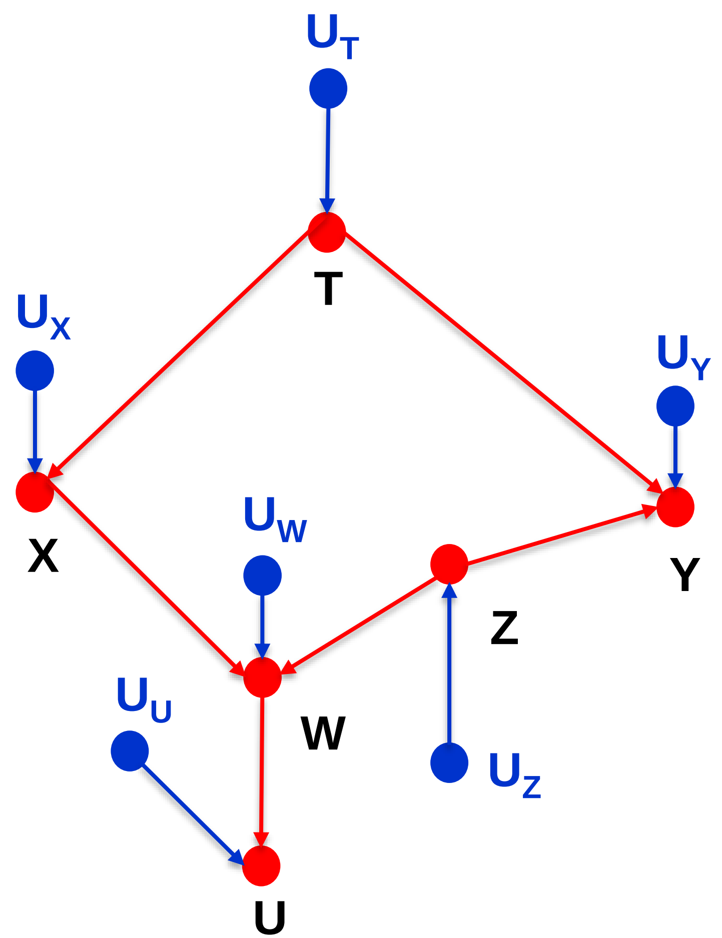
因果模型图与独立性

- **d-分离用于确定因果模型图中任意一对节点是否独立**

- **一条路径会被一组节点 Z 阻断，当且仅当**：
  - 路径 p 包含链结构 $A \to B \to C$ 或分叉结构 $A \leftarrow B \to C$，且中间节点 B 在 Z 中；或
  - 路径 p 包含一个对撞结构 $A \to B \leftarrow C$，且对撞节点 B 及其子孙节点都不在 Z 中。
- **如果一组节点 Z 阻断了 X 和 Y 间的每一条路径，则 X 和 Y 在 Z 的条件下是d-分离的，即$X \perp Y \mid Z$**

» d －分离示例

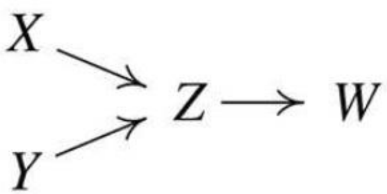
- **以下一组变量是否能形成变量X与Y的 d －分离**
  - {T} √
  - {W} ×
  - {T, U} ×
  - {T, Z} √
  - {W, U, Z} ×
  - {U, Z, T} √



因果模型图示例

应用例子：基于d—分离的因果发现

- 基于 d —分离（独立性）的因果发现
  - 发现与数据对应的因果图模型



真实世界的因果模型（未知）

产生

X_1	Y_1	Z_1	W_1
X_2	Y_2	Z_2	W_2
......			
X_n	Y_n	Z_n	W_n

一堆观测数据（已知）

如何从观测数据中发现因果？

从数据中可以发现以下变量之间的独立性或条件独立性

$X \perp Y$、 $X \not\perp W$、 $X \not\perp Z$、 $Y \not\perp Z$、 $Y \not\perp W$、 $Z \not\perp W$

$Y \not\perp Z|X$、 $Y \not\perp W|X$、 $Z \not\perp W|X$

$X \not\perp W|Y$、 $X \not\perp Z|Y$、 $Z \not\perp W|Y$

$X \not\perp Y|Z$、 $X \perp W|Z$、 $Y \perp W|Z$

$X \perp Y|W$、 $X \not\perp Z|W$、 $Y \not\perp Z|W$

应用例子：基于d—分离的因果发现

- 基于 d —分离（独立性）的因果发现
  - 发现与数据对应的因果图模型

数据中的独立性关系

$X \perp Y$、 $X \not\perp W$、 $X \not\perp Z$、 $Y \not\perp Z$、 $Y \not\perp W$、 $Z \not\perp W$ ①

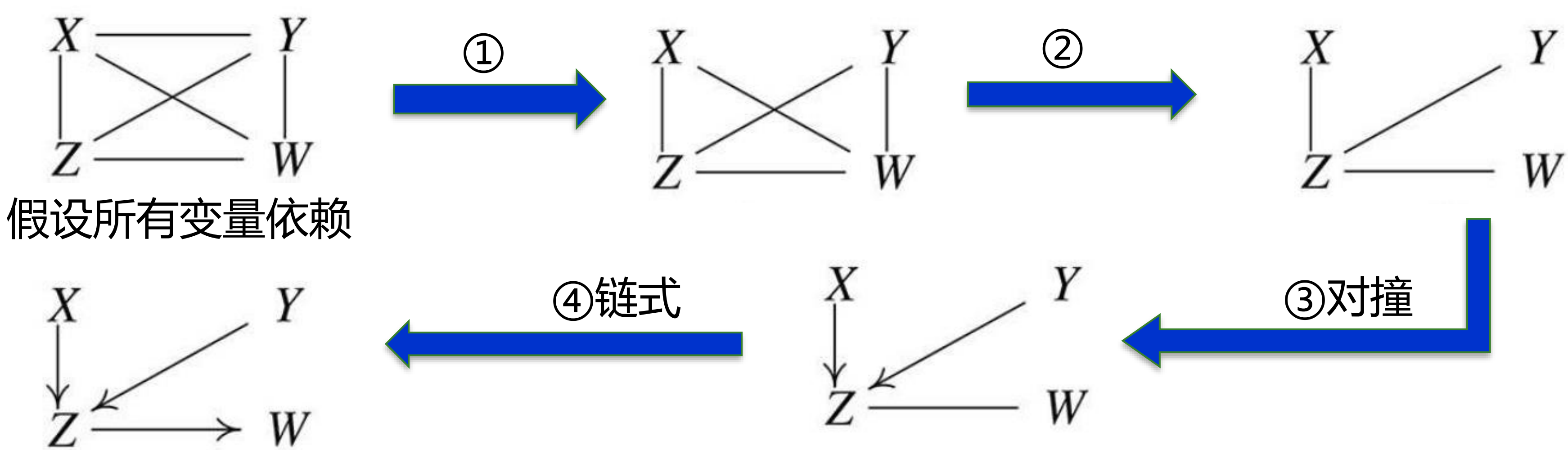
$Y \not\perp Z|X$、 $Y \not\perp W|X$、 $Z \not\perp W|X$

$X \not\perp W|Y$、 $X \not\perp Z|Y$、 $Z \not\perp W|Y$

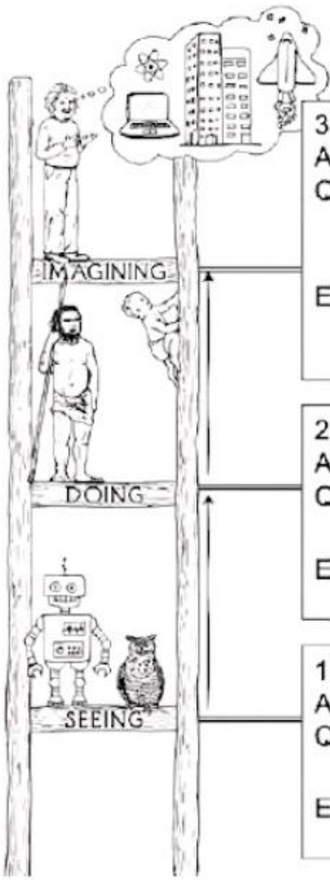
$X \not\perp Y|Z$、 $X \perp W|Z$、 $Y \perp W|Z$

$X \perp Y|W$、 $X \not\perp Z|W$、 $Y \not\perp Z|W$

③ ② ④



因果之梯



3-LEVEL HIERARCHY

3. COUNTERFACTUALS

ACTIVITY: Imagining, Retrospection, Understanding

QUESTIONS: *What if I had done . . . ? Why?*
(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)

EXAMPLES: Was it the aspirin that stopped my headache?
Would Kennedy be alive if Oswald had not killed him? What if I had not smoked the last 2 years?

2. INTERVENTION

ACTIVITY: Doing, Intervening

QUESTIONS: *What if I do . . . ? How?*
(What would Y be if I do X?)

EXAMPLES: If I take aspirin, will my headache be cured?
What if we ban cigarettes?

1. ASSOCIATION

ACTIVITY: Seeing, Observing

QUESTIONS: *What if I see . . . ?*
(How would seeing X change my belief in Y?)

EXAMPLES: What does a symptom tell me about a disease?
What does a survey tell us about the election results?

反事实：回答假设性问题

$$P(Y_{X=x} = y | E = e)$$

干预：研究因果效应

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

联想：数据中寻找相关

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

课间休息

目录

- 基础知识
  - 结构因果模型
  - 因果模型图
- 干预
  - 干预与校正公式
  - 后门准则
  - 前门准则
- 反事实
  - 反事实定义
  - 确定性反事实计算
  - 非确定性反事实计算

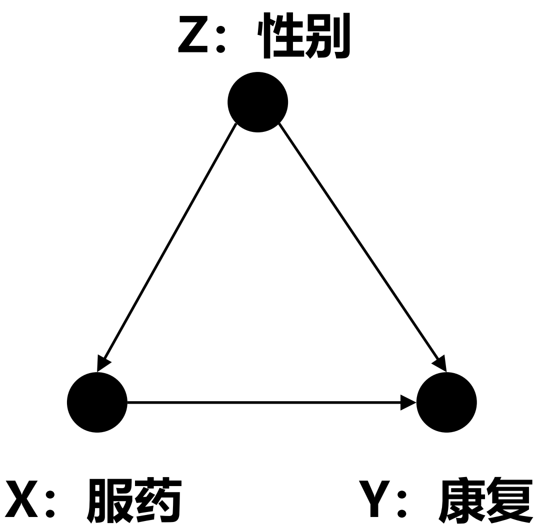
干预效果预测问题

- 很多统计研究问题本质上是预测某种干预措施的效果
  - 通过让患者服用某种新癌症药物（实施干预），病人的病情会如何变化?
  - 减少儿童接触暴力电视节目的干预措施，能否降低儿童的攻击性?
  - 通过控制吸烟，是否能有效降低肺癌的患病率?

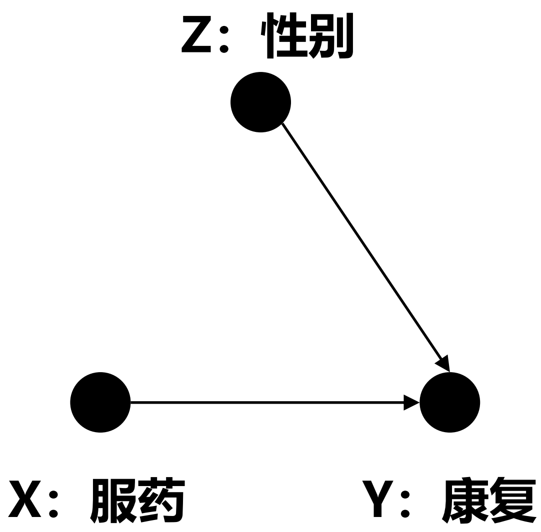
- 这些干预效果预测问题是否能通过因果图来表示?

干预

- **干预（Intervention）的定义**
  - **将变量固定为某个值，限制了该变量随其他变量而变化的自然趋势**
  - **记为$do(X=x)$，在因果模型图中即为去掉所有指向 X 的边**



原因果模型图



干预 X 对应的因果模型图

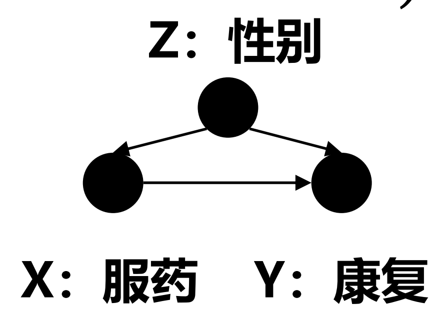
干预 vs. 条件

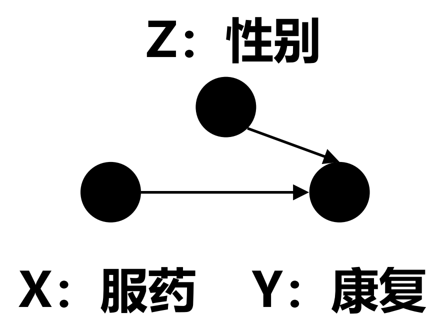
- 对变量进行干预：$P(Y = y|do(X = x))$
  - 通过干预使 $X = x$ 时 $Y = y$ 的概率
  - 如果群体中的每个个体均将变量 X 的值固定为 x 时，Y 的总体分布
  - 改变了世界（系统）本身

- 以变量为条件：$P(Y = y|X = x)$
  - 在 $X = x$ 的条件下 $Y = y$ 的概率
  - 在变量 X 的取值都为 x 的这些个体上，Y 的总体分布
  - 关注问题的子集，改变的仅是我们对世界的看法

干预效果

- **平均因果效应（Average causal effect, ACE）**
  - 以服药（X）是否影响康复（Y）为例，服药对康复的因果效应差异为：
    - $P\big(Y=\text{康复}|do(X=\text{服药})\big)-P\big(Y=\text{康复}|do(X=\text{不服药})\big)$
  - 传统方法：随机对照试验
    - 控制其他变量分布一致，通过改变干预变量，观察其对结果的影响
    - 随机对照试验成本高/有时不可行
  - 能否从观测数据中直接估计干预效果？
    - 校正公式！





校正公式

- 在只有观测数据的情况下，计算干预操作的效果

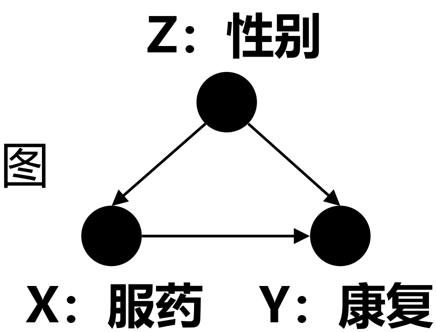
$$P(Y=y|do(X=x))=P_m(Y=y|X=x) \quad \text{(由定义)}$$

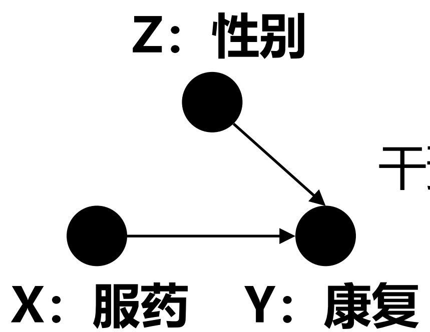
$$=\sum_z P_m(Y=y|X=x,Z=z)P_m(Z=z|X=x) \quad \text{(全概率公式)}$$

$$=\sum_z P_m(Y=y|X=x,Z=z)P_m(Z=z) \quad \text{(条件独立性)}$$

$$=\sum_z P(Y=y|X=x,Z=z)P(Z=z) \quad \text{(函数不变性)}$$

原因果模型图





干预 X 后的因果模型图P_m

校正公式用于解决辛普森悖论

- **校正公式**
  - $P(Y=y|do(X=x))=\sum_z P(Y=y|X=x,Z=z)P\,(Z=z)$

	服药（X=1）	不服药（X=0）
男 (Z=1)	87人中81人康复 (**93%**)	270人234人康复 (**87%**)
女 (Z=0)	263人中192人康复 (**73%**)	80人中55人康复 (**69%**)
合计	350人中273人康复 (**78%**)	350人中289人康复 (**83%**)

男性比例：0.51
(87+270)/700

女性比例：0.49
(263+80)/700

$$P(Y=\text{康复}|do(X=1))=\sum_z P(Y=\text{康复}|X=1,Z=z)P\,(Z=z)=0.93\times0.51+0.73\times0.49=0.832$$

$$P(Y=\text{康复}|do(X=0))=\sum_z P(Y=\text{康复}|X=0,Z=z)P\,(Z=z)=0.87\times0.51+0.69\times0.49=0.7818$$

✓ **服药有效：** $P(Y=\text{康复}|do(X=1))-P(Y=\text{康复}|do(X=0))=0.0502$

一般化的校正公式

因果效应规则

- 给定一个图 G，设变量 X 的父节点集合为 PA，则 X 对 Y 的因果效应为：

$$P(Y=y|do(X=x))=\sum_z P(Y=y|X=x,PA=z)P(PA=z)$$

Proof（同前面的校正公式类似）：

$P(Y=y|do(X=x))=P_m(Y=y|X=x)$ （由定义）

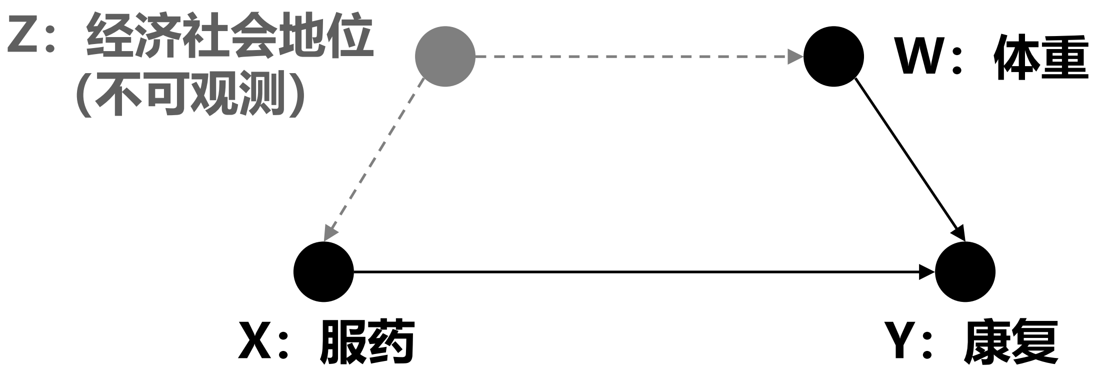
$=\sum_z P_m(Y=y|X=x,PA=z)P_m(PA=z|X=x)$ （全概率公式）

$=\sum_z P_m(Y=y|X=x,PA=z)P_m(PA=z)$ （条件独立性）

$=\sum_z P(Y=y|X=x,PA=z)P\,(PA=z)$ （函数不变性）

如果父节点变量不可观测怎么办?

- 因果模型图示例



- X 的父节点为 Z，那么根据一般化校正公式
  - $P(Y=y|do(X=x))=\sum_z P(Y=y|X=x,Z=z)P(Z=z)$
  - 但由于变量 Z 不可观测，无法从数据中计算上述公式！

» 如果父节点变量不可观测怎么办?

▪ 后门准则

- **给定因果模型图中一对有序变量 (X, Y)，若变量集合 Z 满足：**
  - Z 阻断了 X 与 Y 之间的每条含有指向 X 的路径（后门路径）
  - Z 中没有 X 的后代节点

则称 Z 满足关于(X, Y)的后门准则

▪ 后门校正

- **若变量集合 Z 满足 (X, Y) 的后门准则，那么 X 对 Y 的因果效应可由以下公式计算：**

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

直观理解后门准则

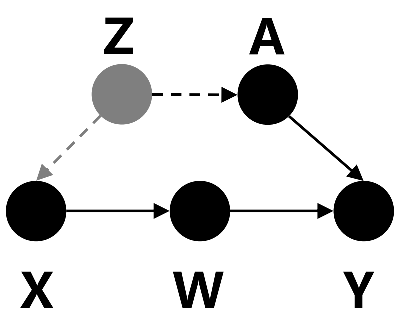
后门准则

- **给定因果模型图中一对有序变量 (X, Y)，若变量集合 Z 满足：**
  - Z 阻断了 X 与 Y 之间的每条含有指向 X 的路径
  - Z 中没有 X 的后代节点

则称 Z 满足关于(X, Y)的后门准则

1. 阻断X和Y之间的后门路径
2. 保持从X到Y的因果路径不变

例子:



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

1. 变量集合希望阻断可能使X和Y存在"伪因果"的后门路径，例如 $X \leftarrow Z \rightarrow A \rightarrow Y$

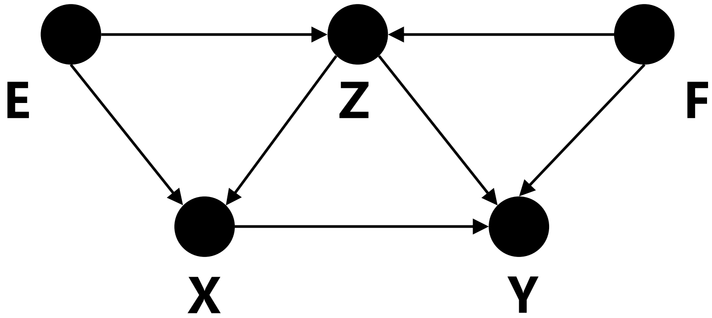
2. 变量W不应该被包含在集合Z中

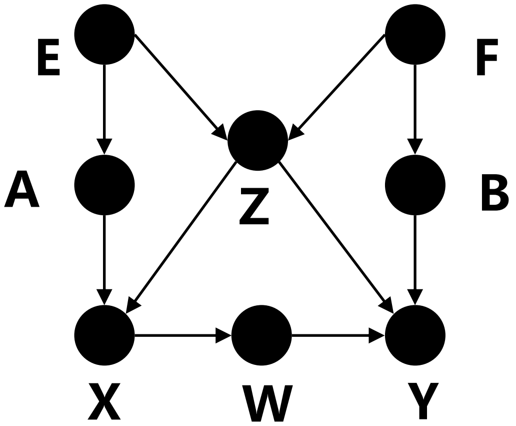
- **一条路径会被一组节点 Z 阻断，当且仅当:**
  - 路径 p 包含链结构 $A \rightarrow B \rightarrow C$ 或分叉结构 $A \leftarrow B \rightarrow C$，且中间节点 B 在 Z 中;
  - 路径 p 包含一个对撞结构 $A \rightarrow B \leftarrow C$，且对撞节点 B 及其子孙节点都不在 Z 中。

该因果模型图中变量 A 满足关于(X, Y)的后门准则

例子：满足后门准则的变量集

- 找出满足变量对 (X, Y) 的后门准则的变量集



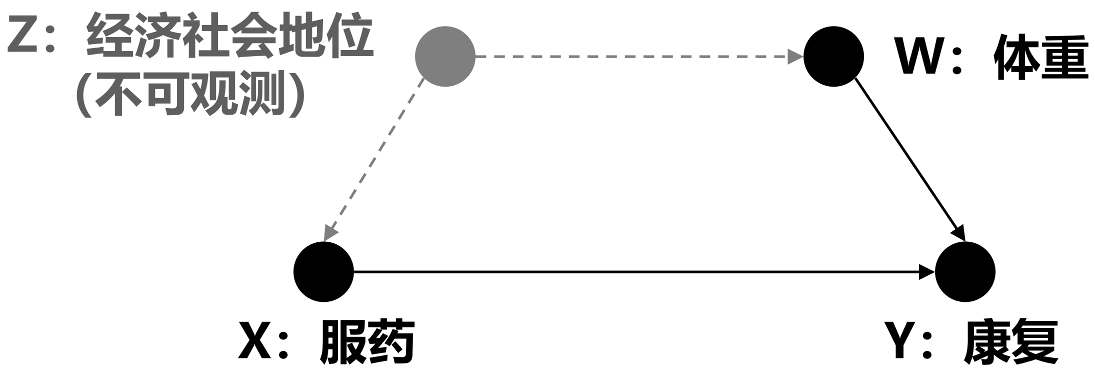


- {Z} ×
- {E, Z} √
- {F, Z} √
- {E, F} ×
- {E, F, Z} √

- {Z} ×
- {F, Z} √
- {A, Z} √
- {W, A, Z} ×
- {A, F, Z} √

例子：应用后门校正估计因果效应

- **因果模型图示例**

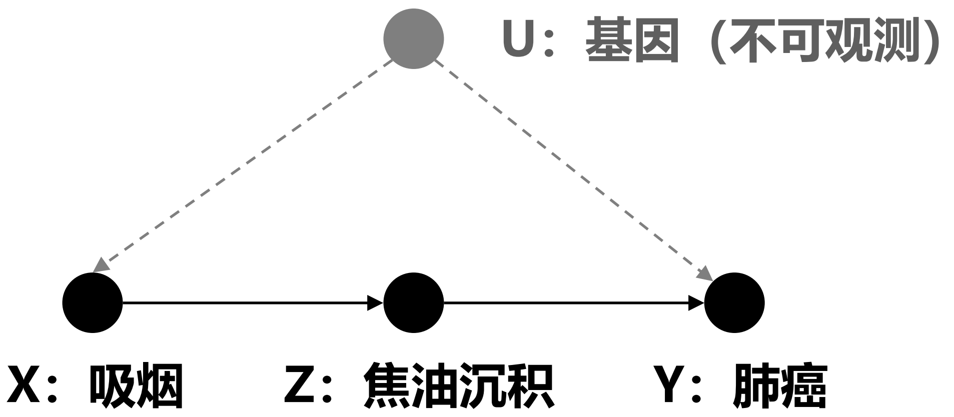


- **变量集合 W 满足变量对 (X, Y) 的后门准则，那么根据后门校正公式：**
  - $P(Y=y|do(X=x)) = \sum_w P(Y=y|X=x, W=w)P(W=w)$

如果后门准则也不满足怎么办?

- 因果模型图示例
  - 需要估计因果效应

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



- 变量 X 的父节点为 U （不可观测）
  - 无法应用一般化校正公式!
- 满足变量对 (X, Y) 后门准则的变量为 U （不可观测）
  - 无法应用后门校正公式!

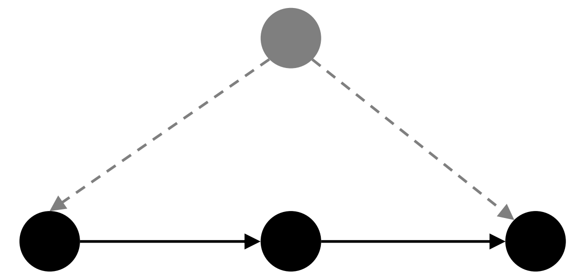
如果后门准则也不满足怎么办?

- **通过中介变量估计因果效应 $P(Y=y|do(X=x))$**
  - **Z 是 X 到 Y 的中介变量**
  - **X 到 Z 的因果效应**

$$P(Z=z|do(X=x))$$
$$=P(Z=z|X=x) \quad \text{(无后门路径)}$$

  - **Z 到 Y 的因果效应**

$$P(Y=y|do(Z=z))$$
$$=\sum_x P(Y=y|Z=z,X=x)P\,(X=x) \quad \text{(后门校正)}$$



- **结合X 到 Z 的因果效应以及Z 到 Y 的因果效应，得到 X 到 Y 的整体因果效应**

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

如果后门准则也不满足怎么办?

前门准则

- **给定因果模型图中一对有序变量 (X,Y)，若变量集合 Z 满足：**
  - Z 切断了所有 X 到 Y 的有向路径;
  - X 到 Z 没有后门路径;
  - 所有 Z 到 Y 的后门路径都被 X 阻断

则称 Z 满足关于(X,Y)的前门准则

前门校正

- **若变量集合 Z 满足 (X,Y) 的前门准则，那么 X 对 Y 的因果效应可由以下公式计算:**

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

目录

- 基础知识
  - 结构因果模型
  - 因果模型图
- 干预
  - 干预与校正公式
  - 后门准则
  - 前门准则
- 反事实
  - 反事实定义
  - 确定性反事实计算
  - 非确定性反事实计算

反事实例子

- 当昨天晚上驾车回家的时候，我在一个路口需要作出选择：走高速公路（$X = 1$）还是走国道（$X = 0$）
- 我选择了走国道（$X = 0$），结果发现交通非常拥堵——用了一个小时（$Y = 1$ 小时）才到家。
- 此时我后悔地想：“如果当初我走高速公路，应该会更快到家吧！”

这种“如果”的陈述形式被称作反事实

反事实内涵与符号表示

- 反事实内涵
  - 反事实陈述句：如果……，那么……
    - “如果”部分称为假设条件
  - 反事实强调**在完全一致的现实条件下，比较不同假设条件的结果**
    - 创造了与真实世界不同的平行世界
- 反事实符号表示
  - 由于假设条件（X）的取值不同而创造出了不同的世界，用下标表示。
  - 不同世界下 Y 的取值记为 $Y_{X=x}$（或 Y_x）

再来看刚才的例子

- **已知现实世界**
  - $X=0$（走国道），$Y=1$（驾驶时间1小时）
- **那么反事实想回答的问题是**
  - 如果当初 $X=1$（走高速），那么驾驶时间 $Y_{X=1}$ 的期望是多少?

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

反事实世界（假设当初走高速）

现实世界（观测变量）

反事实 vs. 干预

- **干预**
  - **讨论的是某个操作所带来的因果效应。没办法区分现实世界和反事实世界。**
    - **例如** $\boldsymbol{P}\left(\text{驾驶时间}|\boldsymbol{do}(\boldsymbol{X}=\text{国道})\right)$，$\boldsymbol{P}\left(\text{驾驶时间}|\boldsymbol{do}(\boldsymbol{X}=\text{高速})\right)$
  - **在总体上进行**

- **反事实**
  - **讨论的是在两个平行世界中发生的不同事件（现实世界与反事实世界），需要保持世界中其他环境/个体条件完全一致**
  - **在个体或群体上进行**

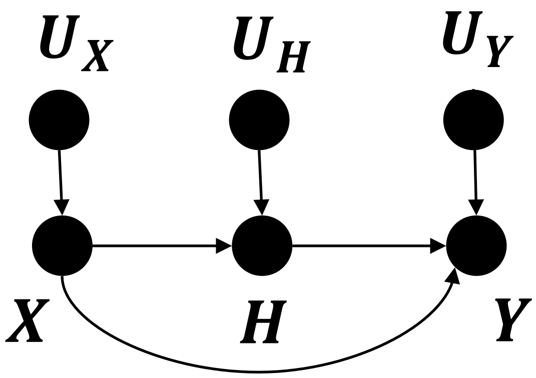
反事实怎么计算?

- 回顾结构因果模型 M
  - 外生变量集合 U，内生变量集合 V，确定内生变量集合的函数/方程集合 $\{F\}$
  - 外生变量取值 $U = u$ 确定了当前个体或环境或情形

- 给定观测$E = e$，反事实 $Y_{X=x}$ 的三步法计算：
  - 溯因：用观测/证据 $E = e$ 确定当前个体/环境，即 U的值
  - 作用：修改结构因果模型 M，移除变量 X 出现在左边的方程并用 $X = x$ 来替换它们，从而获得修正的模型 M_x
  - 预测：使用修正后的模型 M_x 和 U 的值来计算 Y 的值，即反事实结果

例子：课后补习、家庭作业与成绩

- **课后补习、家庭作业量与考试成绩的结构因果模型**
  - X：学生在课后补习上花费的时间
  - H：家庭作业量
  - Y：考生考试成绩



因果模型图

- **考虑一个名为乔的学生**
  - 观测 $X=0.5$，$H=1$，$Y=1.5$
  - 反事实问题：如果家庭作业量 H 加倍，他的考试成绩会更高吗?

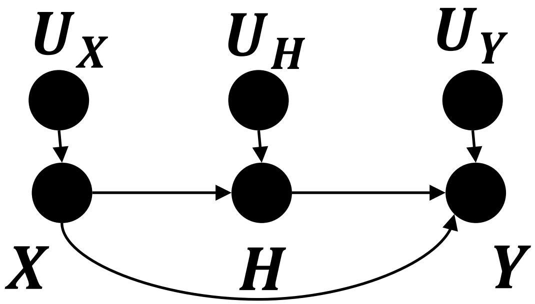
$$X = U_X$$
$$H = 0.5X + U_H$$
$$Y = 0.7X + 0.4H + U_Y$$

结构因果模型

例子：课后补习、家庭作业与成绩

- **反事实计算过程**
  - **溯因：根据 $e=\{X=0.5,\ H=1,\ Y=1.5\}$ 推断个体的 U**
    - $U_X=0.5$，$U_H=1-0.5\times0.5=0.75$，$U_Y=1.5-0.7\times0.5-0.4\times1=0.75$
  - **作用：移除变量 H 出现在左边的方程并用 $H=2$ 来替换**
  - **预测：** $Y_{H=2}(U_x=0.5,U_H=0.75,U_Y=0.75)=1.9>1.5$

反事实答案：如果家庭作业量 ***H*** 加倍，乔的考试成绩会更高！



因果模型图

$$X=U_X$$
$$H=0.5X+U_H$$
$$Y=0.7X+0.4H+U_Y$$

结构因果模型

如果根据观测无法唯一确定个体/环境呢?

- 对于 $E(Y_{X=x}|E=e)$的非确定性反事实，三步法为：

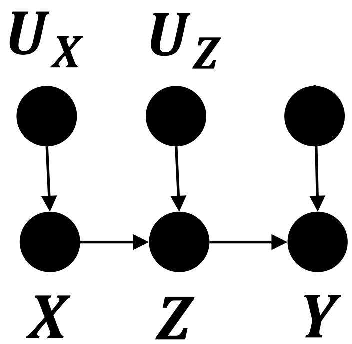
  - **溯因**：用观测/证据 $E=e$ 更新 $P(U)$，获得 $P(U|E=e)$

  - **作用**：修改结构因果模型 M，移除变量 X 出现在左边的方程并用 $X=x$ 来替换它们，从而获得修正的模型 M_x

  - **预测**：使用修正后的模型 M_x 和 $P(U|E=e)$ 的值来计算 Y 的**期望**，即反事实结果

» 例子：教育、技能与薪资

- **教育水平、技能水平与薪资的因果结构模型**
  - X：教育水平
  - Z：工作所需的技能水平
  - Y：薪资



因果模型图

- **对于那些观测到技能水平 $Z = 1$ 的个体**
  - **反事实问题：如果这些个体当初接受了高等教育（$X = 1$），他们的期望薪资会是多少?**

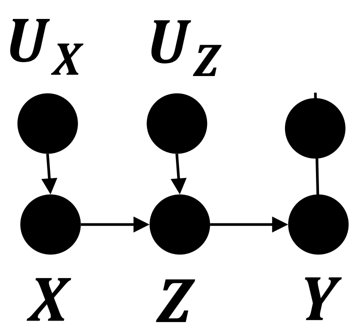
$$X = U_X$$
$$Z = X + U_Z$$
$$Y = Z$$
$$U_X, U_Z \in \{0, 1\}$$

结构因果模型

例子：教育、技能与薪资

- 非确定性反事实计算过程
  - 溯因：用观测 $e=\{Z=1\}$ 更新 $P(U)$，获得 $P(U|E=e)$
    - $P(U_X=0,U_Z=1|Z=1)=\frac{P(U_X=0,U_Z=1)}{P(U_X=0,U_Z=1)+P(U_X=1,U_Z=0)}=p_1$
    - $P(U_X=1,U_Z=0|Z=1)=\frac{P(U_X=1,U_Z=0)}{P(U_X=0,U_Z=1)+P(U_X=1,U_Z=0)}=p_2$
  - 作用：移除变量 X 出现在左边的方程并用 $X=1$ 来替换
  - 预测：$E(Y_{X=1}|Z=1)=p_1Y_{X=1}(U_X=0,U_Z=1)+p_2Y_{X=1}(U_X=1,U_Z=0)=2p_1+p_2$

因果模型图



$$
\begin{gathered}
X=U_X \\
Z=X+U_Z \\
Y=Z \\
U_X,U_Z\in\{0,1\}
\end{gathered}
$$

结构因果模型

» 如果未知因果结构模型怎么办?

▪ 后门的反事实定理

- 如果一个变量集合 Z 满足 (X, Y) 的后门条件，那么对于所有的 x，在给定 Z 条件下，反事实条件 $Y_{X=x}$ 独立于 X，记为:

$$P(Y_x|X,Z) = P(Y_x|Z)$$

▪ 反事实结果可以通过观测数据进行估计:

$$P(Y_x = y) = \sum_z P(Y_x = y|Z = z)P(z) \quad \text{(全概率公式)}$$

$$= \sum_z P(Y_x = y|Z = z, X = x)P(z) \quad \text{(后门的反事实定理)}$$

$$= \sum_z P(Y = y|Z = z, X = x)P(z) \quad \text{(反事实一致性原则)}$$

小结

- 统计中的因果推断具有重要意义
  - 克服基于统计相关的机器学习面临的一些问题
    - 模型鲁棒性问题：OOD（out of distribution）
    - 反事实问题：模型去偏等
- 因果模型图给因果推断提供了简洁的数学工具
  - 独立性判定
  - 干预效应
  - 反事实推理
- 因果推断是近年来的学术前沿，其应用方兴未艾

下课