人工智能实践
Artificial Intelligence Practice

DCS3015 Autumn 2022

Chao Yu （余超）

School of Computer Science and Engineering
Sun Yat-Sen University

Lecture 3：因果推断（Causal Inference）

因果推断

因果推断有什么用？

推断某些政策、干预带来的效应，帮助人们进行决策。

一些例子：

- 研究治疗方案对疾病的治疗效果
- 研究环保政策对气候的影响
- 研究社交媒体对心理健康的影响
- …

辛普森悖论

为什么要学习因果推断？——从辛普森悖论开始说起

假定现在存在一种疾病，
现有两种治疗方案（T）：A（0），B（1）
存在两种病情（C）：轻微（0），严重（1）
治疗结果（Y）：存活（0），死亡（1）

辛普森悖论

经过治疗后，得到以下的死亡率数据：

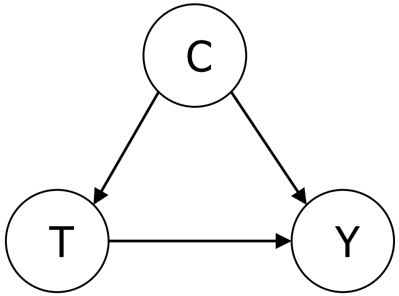
	轻微	严重	总体
方案A	15% (210/1400)	30% (30/100)	16% (240/1500)
方案B	10% (5/50)	20% (100/500)	19% (105/550)

仅看整体数据，治疗方案A死亡率更低，我们应该继续使用方案A。可是如果我们从两种病情分别去看，似乎方案B更加有效？
也就是说如果我们知道病人的病情，我们应该使用方案B，如果不知道则使用方案A。显然，我们得出了一个相悖的结论！

辛普森悖论

是什么影响了我们的推断？——混淆变量（Confounder）

	轻微	严重	总体
方案A	15% (210/1400)	30% (30/100)	16% (240/1500)
方案B	10% (5/50)	20% (100/500)	19% (105/550)



因果图（Causal Graph）

一种可能的因果关系是病情（C）同时影响治疗方案（T）和治疗结果（Y）：

- 病情轻微：倾向使用方案A且死亡率低
- 病情严重：倾向使用方案B且死亡率高

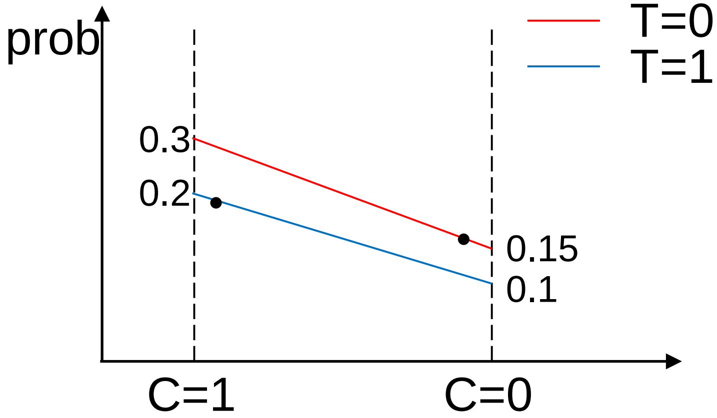
我们称这里的病情为**混淆变量**。为了测量治疗方案的真实效果，必须排除混淆变量的干扰。

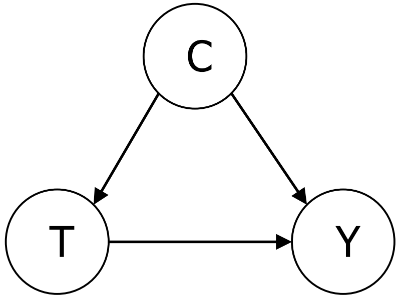
辛普森悖论

为什么在不同病情中方案B都更优，整体却是方案A更优？

$$P(Y=1|T=0)=\sum_{c=0,1}P(Y=1,C=c|T=0)$$
$$=P(Y=1|C=0,T=0)P(C=0|T=0)$$
$$+P(Y=1|C=1,T=0)P(C=1|T=0)$$
$$=0.15*\boxed{\frac{1400}{1500}}+0.3*\boxed{\frac{100}{1500}}=0.16$$

$$P(Y=1|T=1)=\sum_{c=0,1}P(Y=1,C=c|T=1)$$
$$=P(Y=1|C=0,T=1)P(C=0|T=1)$$
$$+P(Y=1|C=1,T=1)P(C=1|T=1)$$
$$=0.1*\boxed{\frac{50}{550}}+0.2*\boxed{\frac{105}{550}}=0.19$$





因果图（Causal Graph）

	轻微	严重	总体
方案A	15% (210/1400)	30% (30/100)	16% (240/1500)
方案B	10% (5/50)	20% (100/500)	19% (105/550)

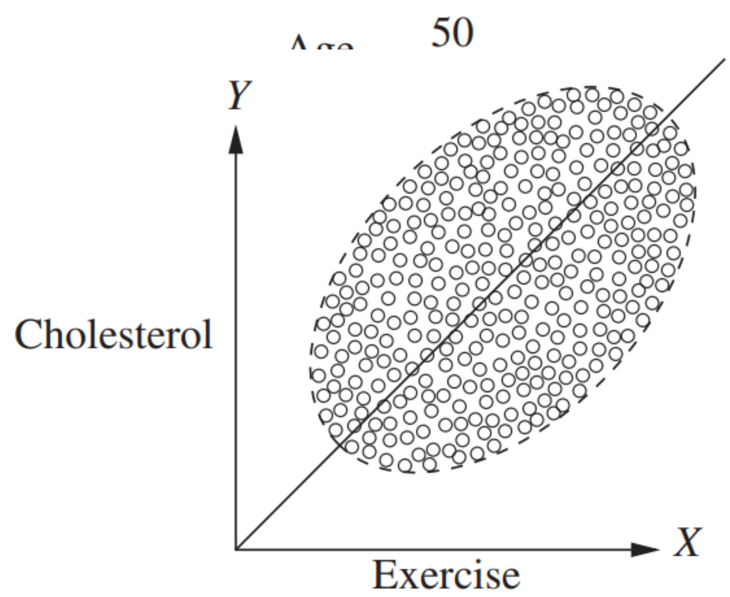
可以看到这里的加权系数对方案B来说是不公平的，为什么会产生这样的加权系数？

辛普森悖论——另一个案例

现有一份不同年龄段人群的锻炼量与胆固醇含量关系的统计数据

在不同的年龄段中，我们都可以看到锻炼量与胆固醇含量之间明显的负相关关系。

可当我们取消年龄这个维度再来看数据，却得到了锻炼量与胆固醇含量之间的正相关关系。进而导致我们得到错误的推论：锻炼的越多，胆固醇含量越高。



辛普森悖论

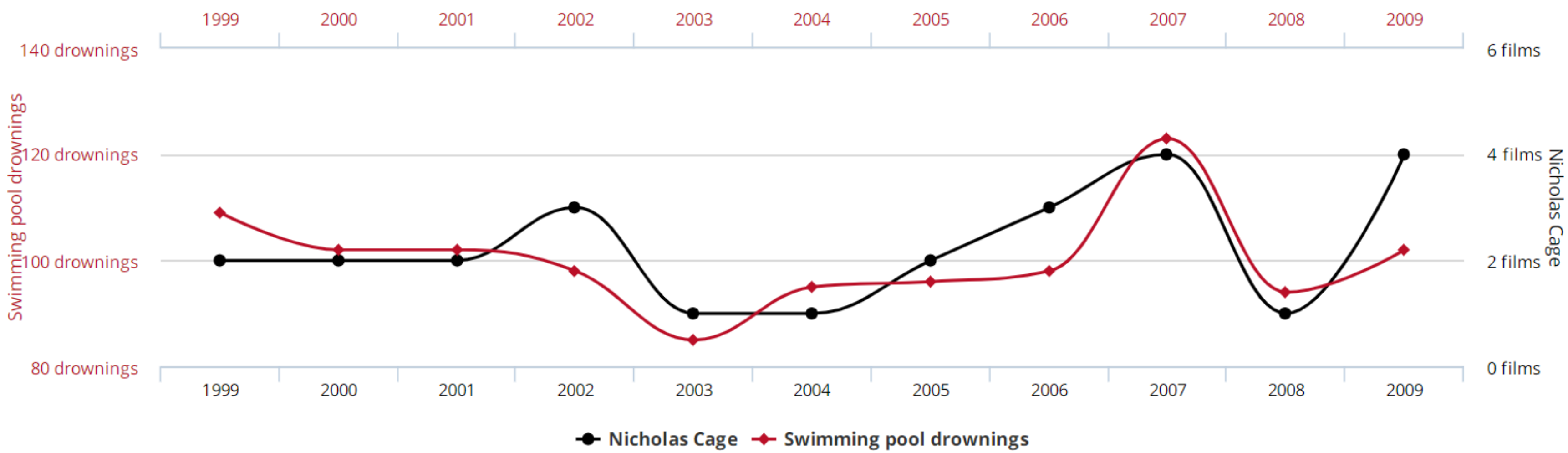
观察以下案例，思考是否需要分类数据来进行决策。

假定现对于肾结石存在两种治疗方案：方案A和方案B。而医生倾向于对较大结石（病情更为严重）使用治疗方案A，对小结石使用方案B。那么如果一个病人尚不知道自己的结石大小，他应该观察整体的治疗数据还是按结石大小分类后的治疗数据呢？（治疗数据即不同方案的治愈率数据）

相关性与因果性

相关性不意味着因果性

下图数据显示每年游泳溺水人数与尼古拉斯·凯奇每年出演的电影数目呈现出高度的统计相关性[1]。所以这意味着尼古拉斯·凯奇观察到溺水人数增多，遂有意参演更多电影来宣扬相关安全问题？又或者是他有兴趣增加他在因果推断从业者中的知名度，所以刻意人为的制造了这些数据？显然这些推断是荒谬的，以上两个数据之间的统计相关性是偶然的，因此也不具备因果关系。

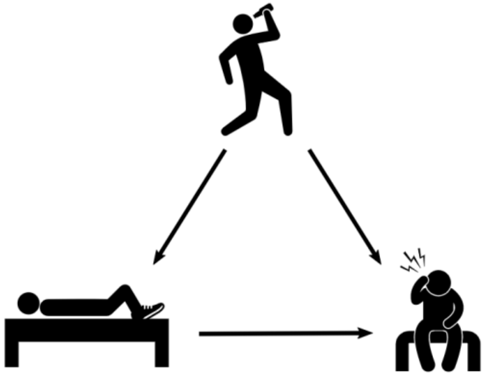


相关性与因果性

相关性不意味着因果性

假设你现有一笔研究穿鞋睡觉与醒来头疼之间关系的数据。数据显示绝大多数时候人们穿鞋睡觉都会在醒来时头疼；而人们不穿鞋睡觉时，则一般不会在醒来时觉得头疼。
人们在遇到这样的数据时，通常会将二者联系起来并得出以下的结论：穿鞋睡觉会导致醒来时头疼。一些小编甚至进而据此结论编写这样的文章："长期穿鞋睡觉会增大阿尔兹海默症患病概率！"

而事实上我们知道，这两者都可能受一个共同原因影响：睡前大量饮酒。因此我们观察到的相关性实际由两部分组成：1）饮酒带来的相关关系；2）穿鞋睡觉与醒来头疼之间的因果关系。而第二部分才是我们在寻找的因果关系。



结构化因果模型

为了进一步研究因果性，我们引入结构化因果模型（Structural Causal Models，SCM）来形式化描述不同元素之间的相关关系，进而研究他们是如何相互影响的。

一个SCM包括两个变量集合，U和V，以及一个函数集合F。

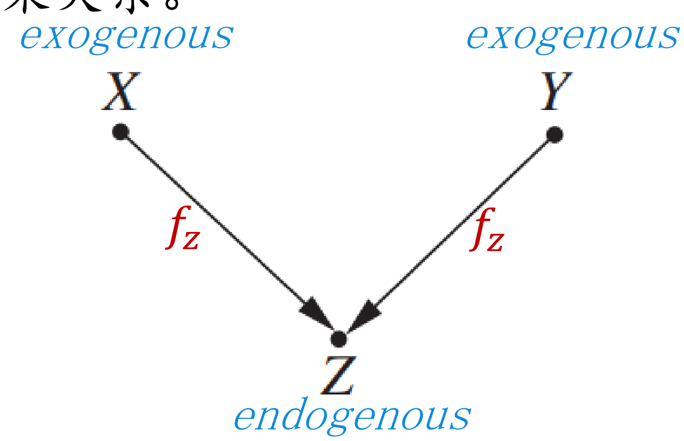
- U：U中的变量称为外生变量（exogenous variables），我们引入这些变量，但我们不考虑这些变量如何被其它变量影响。
- V：V中的变量称为内生变量（endogenous variables），我们想知道它们与其他变量之间的因果关系。
- F：集合F中的元素是一系列函数，这些函数根据其他变量来决定集合V中变量的取值。

每个SCM都存在一个对应的图，U、V中的变量对应图中的节点，F中的函数即对应图中的边。如果以图论语言描述，那么外生变量与内生变量存在以下关系：每一个内生变量是至少一个外生变量的后代；外生变量不能是模型中其它变量的后代，即外生变量为图中的根节点。

结构化因果模型——示例一：薪酬水平

$$U=\{X,Y\},V=\{Z\},F=\{f_Z\}$$
$$f_Z:Z=2X+3Y$$

以上SCM模型表示薪酬（Z）与受教育年限（X）以及工作年限（Y）之间的关系。由于X和Y同时出现在了函数f_Z中，因此称X和Y是Z的直接原因（direct cause）。如果X和Y还存在祖先，那么称这些祖先是Z的潜在原因（potential cause）。该SCM对应的图如下所示，通过该图可以更直观的看出变量之间的因果关系。

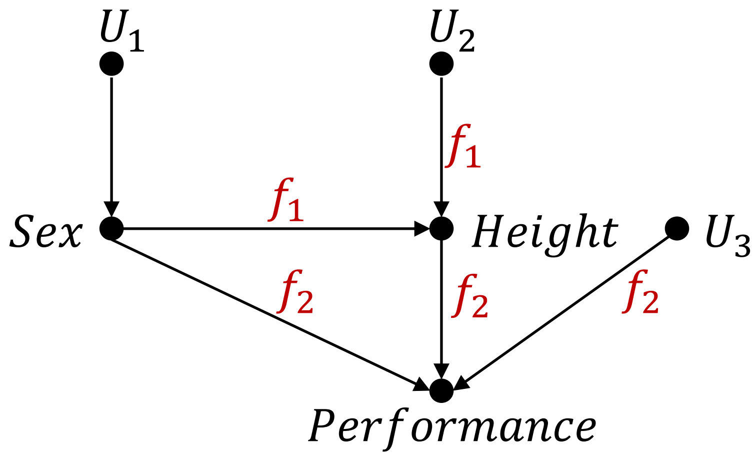


结构化因果模型

结构化因果模型——示例二：篮球表现

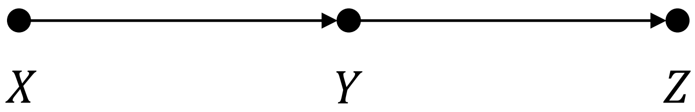
$U = \{U_1, U_2, U_3\}, V = \{Height, Sex, Performance\}, F = \{f_1, f_2\}$
$Sex = U_1,$
$Height = f_1(Sex, U_2),$
$Performance = f_2(Height, Sex, U_3)$

这里的$U = \{U_1, U_2, U_3\}$为我们并不关心如何命名且也不易测量的外生变量，但是它们会影响我们关心的V中的内生变量。尝试画出该SCM对应的图模型？



结构化因果模型

Intransitive Case



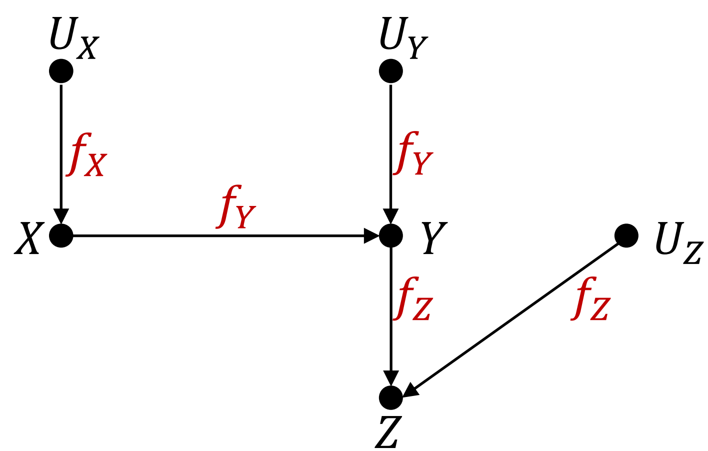
X是Y的直接原因，Y是Z的直接原因，但X可能与Z独立。

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

f_X: $X = U_X$,

$$f_Y\colon\ Y = \begin{cases} a \ if\ X = 1, U_Y = 1 \\ b \ if\ X = 2, U_Y = 1 \\ c \ if\ U_Y = 2 \end{cases}$$

$$f_Z\colon\ Z = \begin{cases} i \ if\ Y = c\ or\ U_Z = 1 \\ j \ if\ U_Z = 1 \end{cases}$$

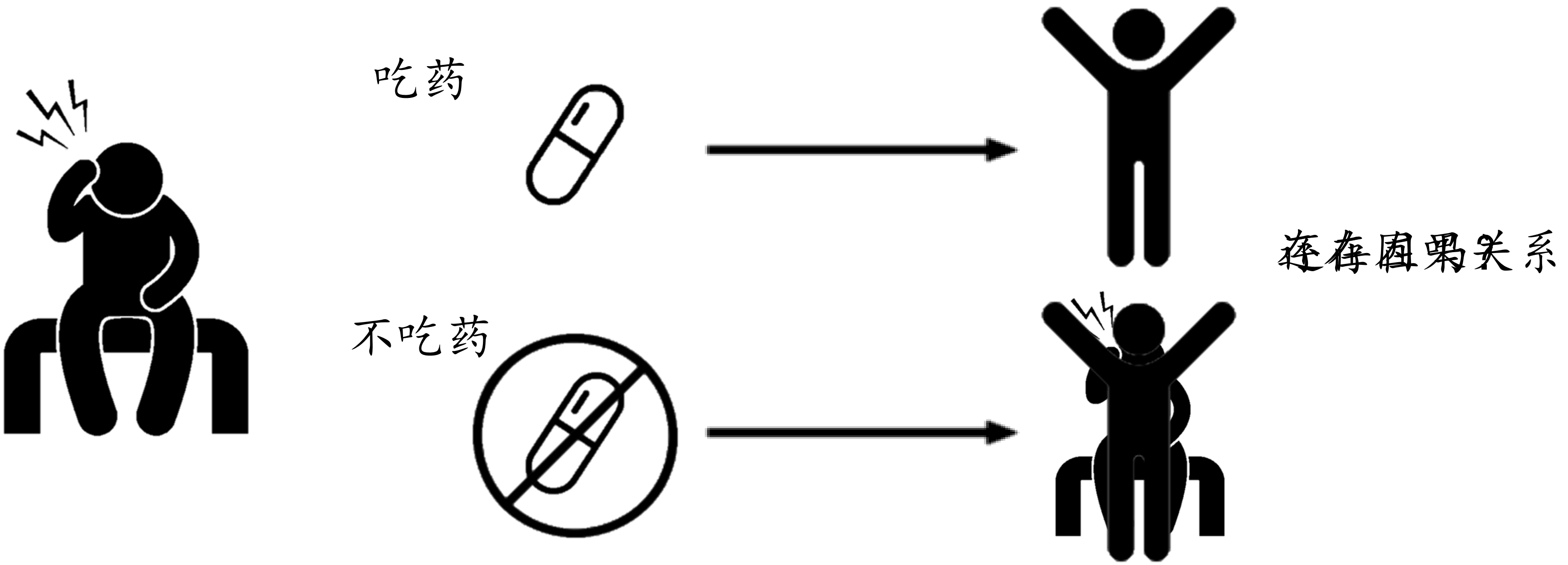


在这个例子中，改变X的取值只会使得Y取值a或b，而仅当Y取值为c时才会影响Z的取值，因此X和Z是独立的。

潜在结果

潜在结果（Potential Outcome）

因果推断——推断动作或策略对某些结果的影响

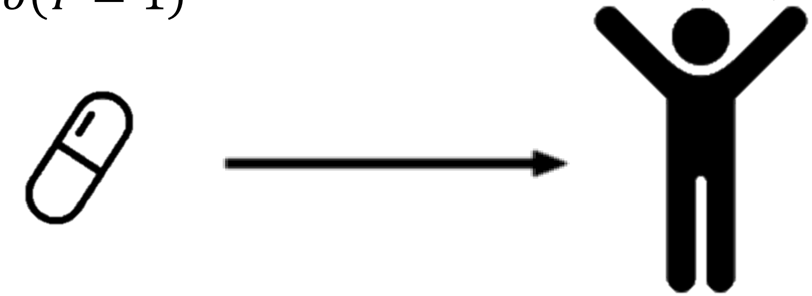


潜在结果

潜在结果：notation

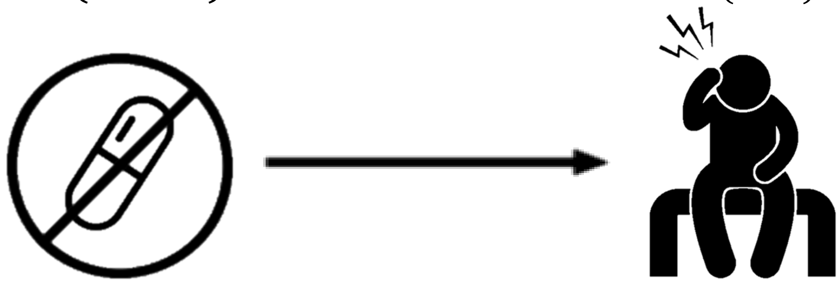
$do(T=1)$

$Y_i|_{do(T=1)} \triangleq Y_i(1)$



$do(T=0)$

$Y_i|_{do(T=0)} \triangleq Y_i(0)$



T：实施的动作

Y：观察到的结果

i：作为下标表示具体的个体

$Y_i(1)$：实施动作的潜在结果

$Y_i(0)$：不实施动作的潜在结果

Causal effect

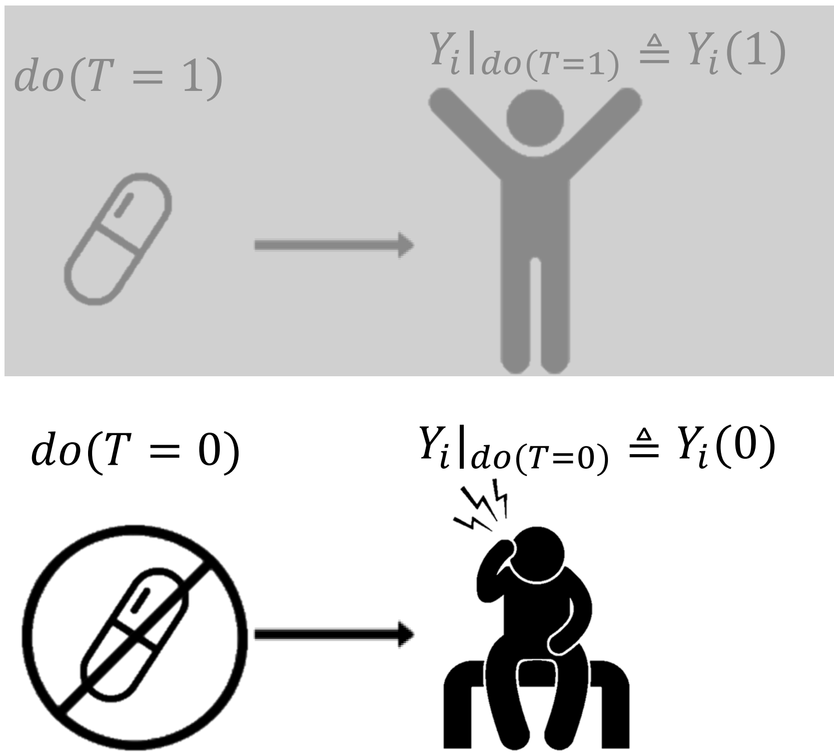
$Y_i(1) - Y_i(0)$

因果推断的基本问题

因果推断的基本问题

Counterfactual

Factual



T：实施的动作

Y：观察到的结果

i：作为下标表示具体的个体

$Y_i(0)$：实施动作的潜在结果

$Y_i(0)$：不实施动作的潜在结果

Causal effect

$Y_i(1) - Y_i(0)$

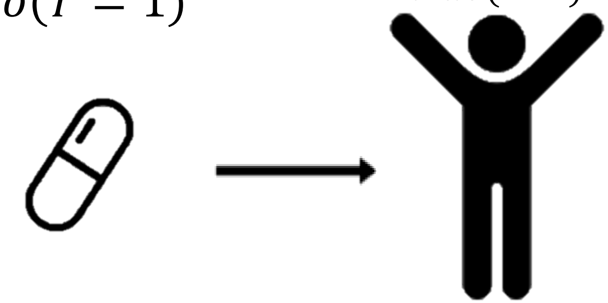
因果推断的基本问题

因果推断的基本问题

Factual

$do(T=1)$

$Y_i|_{do(T=1)} \triangleq Y_i(1)$



Counterfactual

$do(T=0)$

$Y_i|_{do(T=0)} \triangleq Y_i(0)$

T：实施的动作

Y：观察到的结果

i：作为下标表示具体的个体

$Y_i(0)$：实施动作的潜在结果

$Y_i(0)$：不实施动作的潜在结果

Causal effect

$Y_i(1) - Y_i(0)$

因果推断的基本问题

平均处理效应(Average Treatment Effect, ATE)

$$\mathbb{E}[Y(1)-Y(0)]=\mathbb{E}[Y(1)]-\mathbb{E}[Y(0)] \neq \underline{\mathbb{E}[Y|T=1]}-\underline{\mathbb{E}[Y|T=0]}$$

i	T	Y	Y(1)	Y(0)	Y(1)-Y(0)
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	1	?	?

$^2/_3$ $^1/_3$

$$^2/_3-{}^1/_3={}^1/_3$$

T：实施的动作

Y：观察到的结果

i：作为下标表示具体的个体

$Y_i(0)$：实施动作的潜在结果

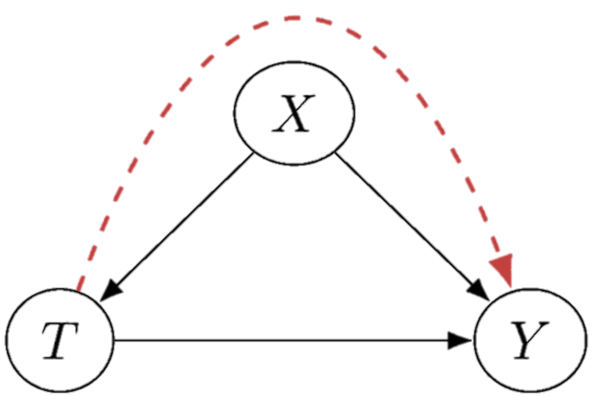
$Y_i(0)$：不实施动作的潜在结果

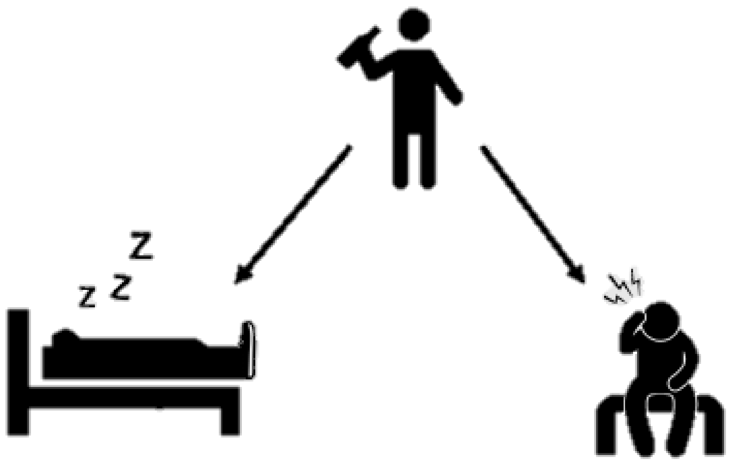
因果推断的基本问题

相关性不意味着因果性

例如穿鞋睡觉和醒来头疼有着很强的相关性，但这源于它们有一个共同原因：睡前喝了大量的酒。

因此穿鞋睡觉的群体和不穿鞋睡觉的群体除了是否穿鞋不同以外，还受这样一个重要变量的影响。



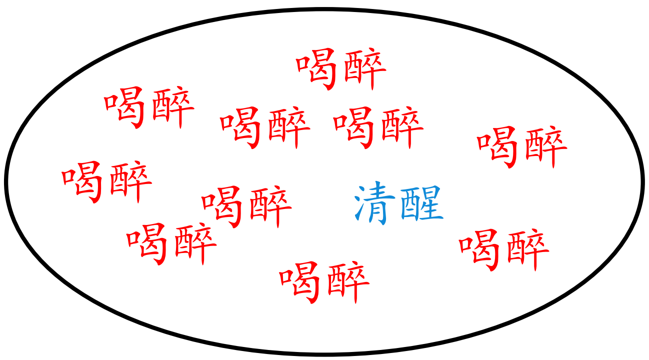


因果推断的基本问题

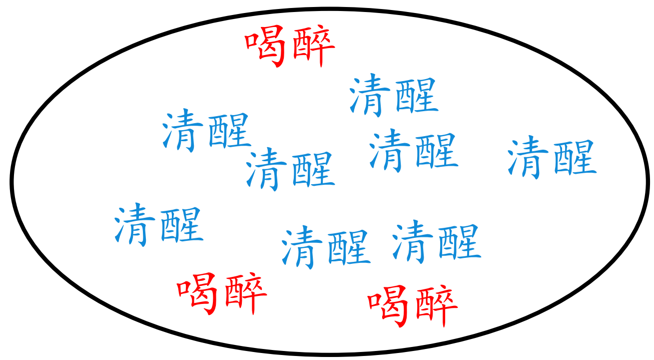
这两个群体不能直接比较，我们需要控制变量

$$\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \neq \mathbb{E}[Y|T=1] - \mathbb{E}[Y|T=0]$$

穿鞋睡觉的群体$(T=1)$



不穿鞋睡觉的群体$(T=0)$

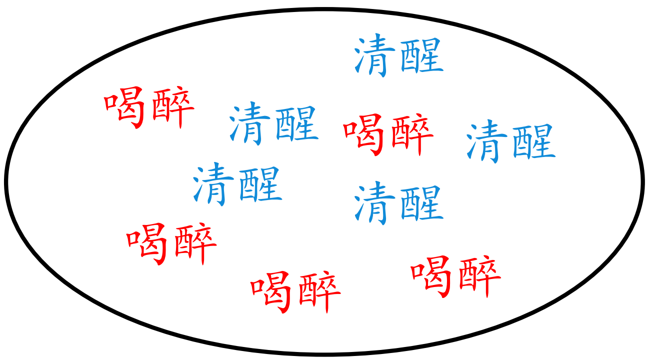


因果推断的基本问题

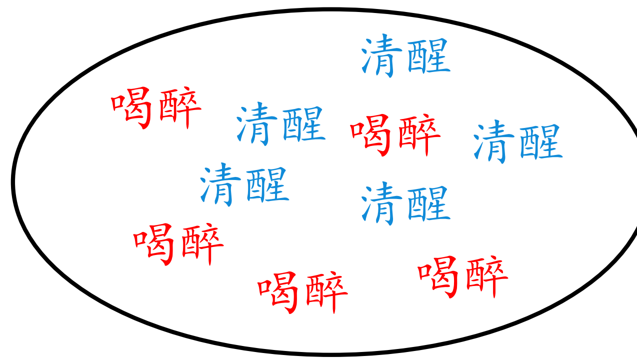
那么什么样的群体是可以比较的？

$$\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y|T=1] - \mathbb{E}[Y|T=0]$$

穿鞋睡觉的群体$(T=1)$



不穿鞋睡觉的群体$(T=0)$



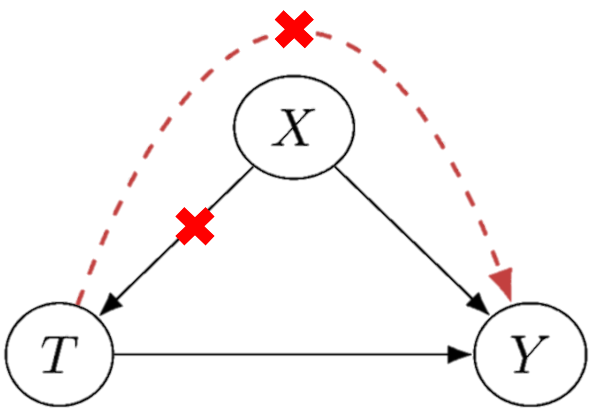
因果推断的基本问题

什么条件下我们可以认为相关性等于因果性？

因果推断的基本问题

Ignorability: $(\boldsymbol{Y}(\mathbf{1}), \boldsymbol{Y}(\mathbf{0})) \perp\!\!\!\perp \boldsymbol{T}$

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)|T] - \mathbb{E}[Y(0)|T]$$
$$= \mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0]$$



另一种角度：exchangeability

$T = 1$ $T = 0$

$\mathbb{E}[Y|T = 1] = y_1$ $\mathbb{E}[Y|T = 0] = y_0$

因果推断的基本问题

另一种角度：exchangeability

$T = 1$ $\qquad\qquad$ $T = 0$

$\mathbb{E}[Y|T = 1] = y_1$ $\qquad\qquad$ $\mathbb{E}[Y|T = 0] = y_0$

交换前 交换后

$$\mathbb{E}[Y(1)|T = 1] = \mathbb{E}[Y(1)|T = 0] = \mathbb{E}[Y(1)]$$
$$\mathbb{E}[Y(0)|T = 0] = \mathbb{E}[Y(0)|T = 1] = \mathbb{E}[Y(0)]$$

因果推断的基本问题

Identifiability

$$\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1)|T=1] - \mathbb{E}[Y(0)|T=0] (ignorability)$$

因果性度量 $= \mathbb{E}[Y|T=1] - \mathbb{E}[Y|T=0]$

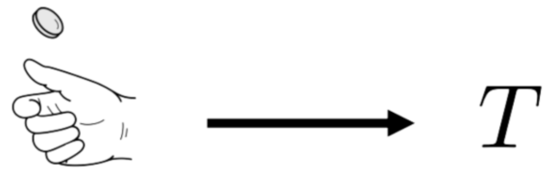
相关性度量

（如果我们知道 $P(x,t,y)$，那我们可以计算得到该项数据）

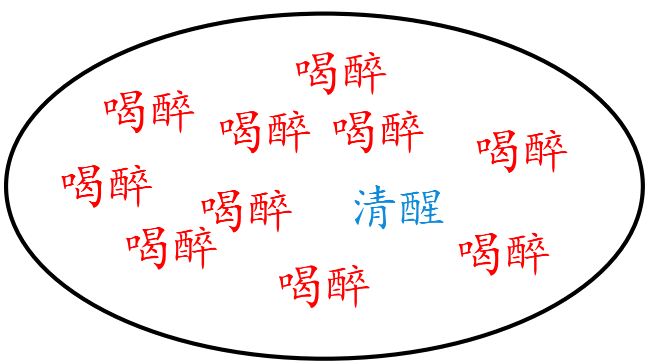
如果一个因果性度量（如$\mathbb{E}[Y(1)]$）可以利用某些统计数据（如$\mathbb{E}[Y|\mathrm{t}]$）计算得出，那么称该因果性度量是**可识别的（identifiability）**。

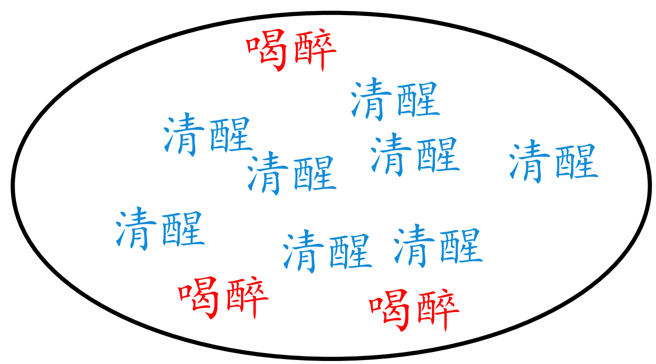
因果推断的基本问题

随机对照试验(Randomized control trial, RCT)



穿鞋睡觉的群体($T = 1$)　　不穿鞋睡觉的群体($T = 0$)





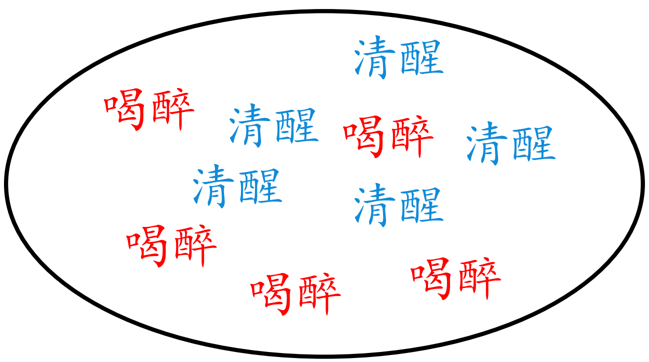
随机对照试验的基本方法是，将研究对象**随机分组**，对不同组实施不同的**干预**，在这种严格的条件下对照效果的不同。在研究对象数量足够的情况下，这种方法可以抵消已知和未知的混杂因素对各组的影响。

因果推断的基本问题

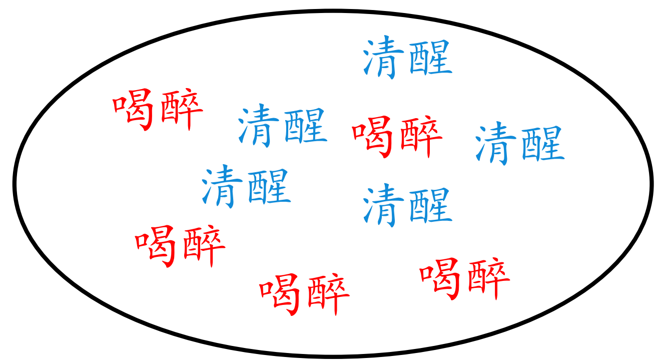
随机对照试验(Randomized control trial, RCT)

$\longrightarrow T$

穿鞋睡觉的群体$(T=1)$



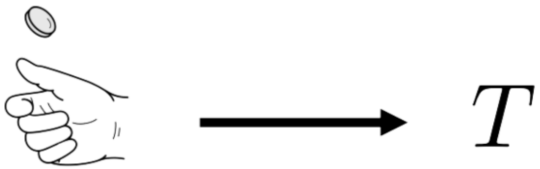
不穿鞋睡觉的群体$(T=0)$

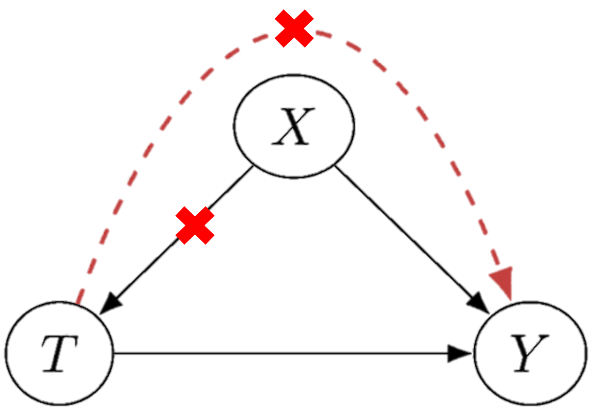


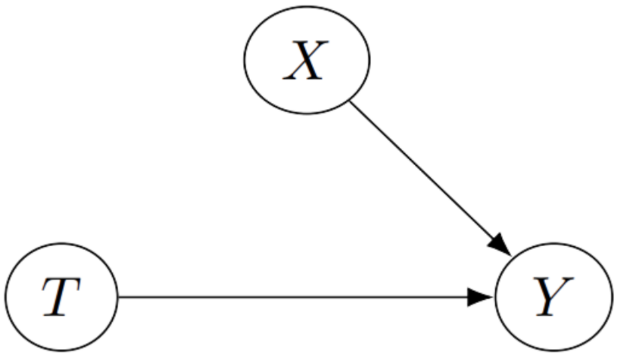
随机对照试验的基本方法是，将研究对象**随机分组**，对不同组实施不同的**干预**，在这种严格的条件下对照效果的不同。在研究对象数量足够的情况下，这种方法可以抵消已知和未知的混杂因素对各组的影响。

因果推断的基本问题

RCT的图形化阐释







因果推断的基本问题

但并不总是能使用对照试验

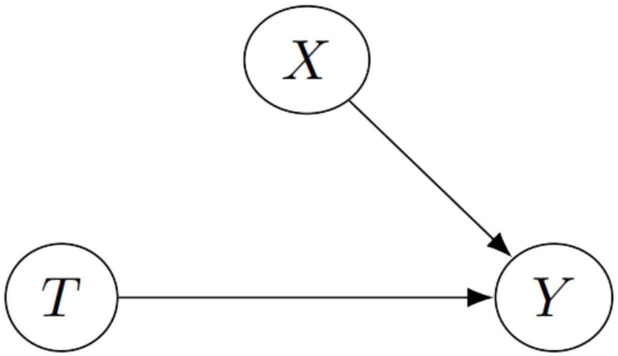
- 道德原因：比如不能随机让人们抽烟来检验其对肺癌的影响
- 方案不可行：比如不能随机让不同国家实施共产主义/资本主义制度来测量制度对GDP的影响
- 完全不可能：比如不能改变一个人的DNA来测量其对身高的影响

因果推断的基本问题

Conditional exchangeability

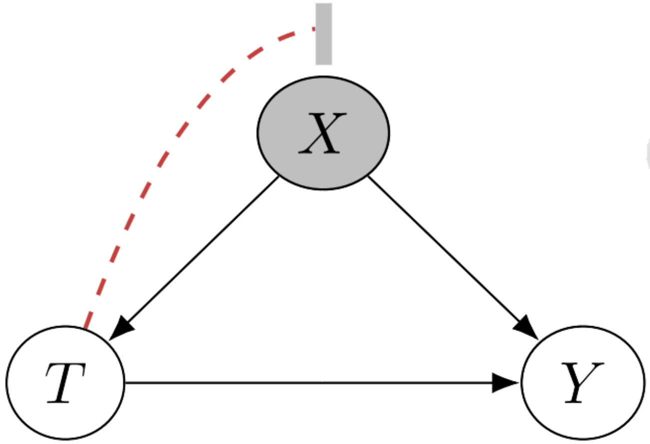
Exchangeability

$(Y(1), Y(0)) \perp\!\!\!\perp T$



Conditional exchangeability:

$(Y(1), Y(0)) \perp\!\!\!\perp T | \mathrm{X}$



因果推断的基本问题

识别条件平均处理效应(Identification of conditional average treatment effect)

$$\begin{aligned}\mathbb{E}[Y(1)-Y(0)|X] &= \mathbb{E}[Y(1)|X]-\mathbb{E}[Y(0)|X] \\ &= \mathbb{E}[Y(1)|T=1,X]-\mathbb{E}[Y(0)|T=0,X] \\ &= \mathbb{E}[Y|T=1,X]-\mathbb{E}[Y|T=0,X]\end{aligned}$$

但我们想要知道的是$\mathbb{E}[Y(1)-Y(0)]$，怎么计算？

调整公式(The Adjustment Formula)

$$
\begin{aligned}
\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) | X] \\
&= \mathbb{E}_X[\mathbb{E}[Y | T = 1, X] - \mathbb{E}[Y | T = 0, X]]
\end{aligned}
$$

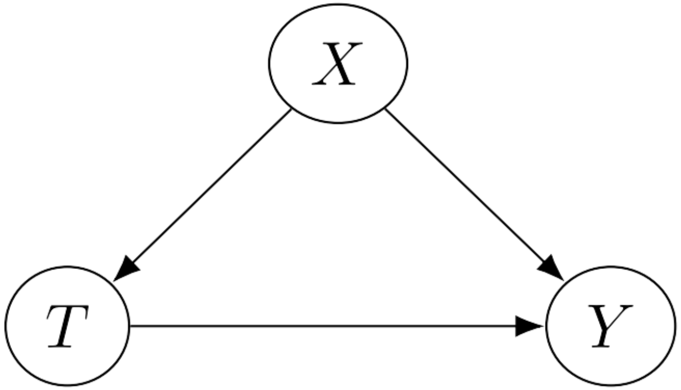
因果推断的基本问题

Unconfoundedness假设往往不易验证

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

$(Y(1), Y(0)) \perp\!\!\!\perp T | X$



因果推断的基本问题

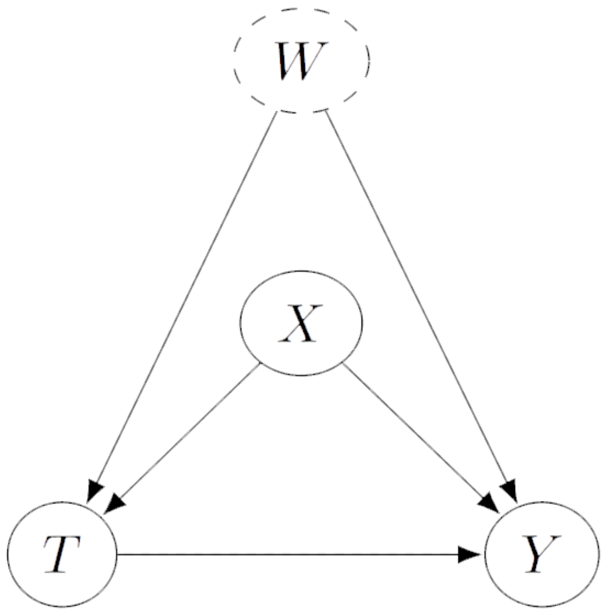
Unconfoundedness假设往往不易验证

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

$$(Y(1), Y(0)) \not\perp\!\!\!\perp T \mid X$$

但如果存在不知道的混淆变量以上假设就不再成立了



因果推断的基本问题

Positivity

对于所有需要控制的协变量x，满足以下条件：

$$0 < P(T=1|X=x) < 1$$

回忆调整公式：

$$\begin{aligned}\mathbb{E}[Y(1)-Y(0)] &= \mathbb{E}_X\mathbb{E}[Y(1)-Y(0)|X] \\ &= \mathbb{E}_X[\mathbb{E}[Y|T=1,X]-\mathbb{E}[Y|T=0,X]]\end{aligned}$$

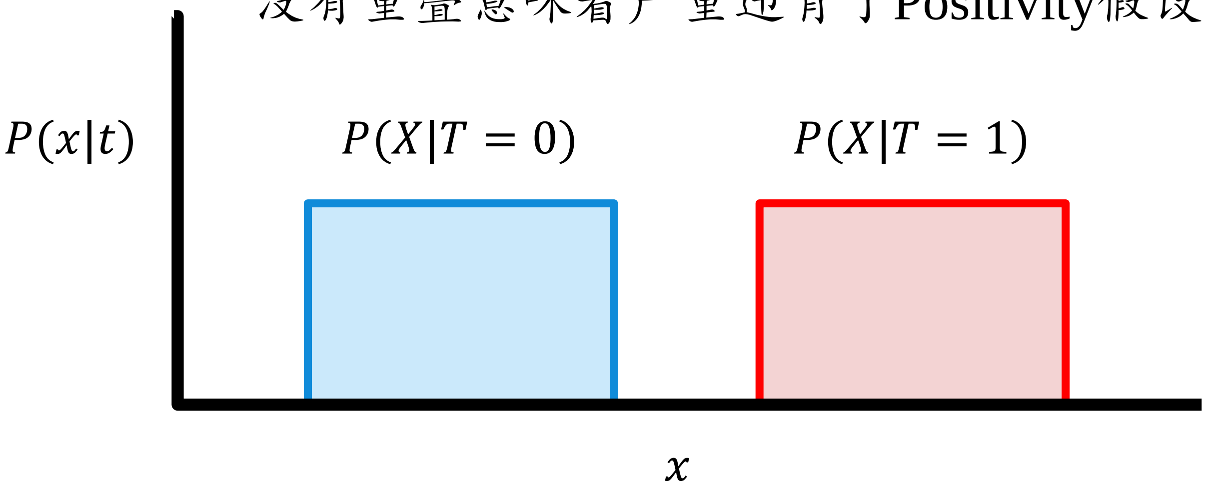
$$\sum_x(\sum_y yP(Y=y|T=1,X=x)-\sum_y yP(Y=y|T=0,X=x))$$

$$\sum_x(\sum_y y\frac{P(Y=y|T=1,X=x)}{\underline{P(T=1|X=x)}P(X=x)}-\sum_y y\frac{P(Y=y|T=0,X=x)}{\underline{P(T=0|X=x)}P(X=x)})$$

因果推断的基本问题

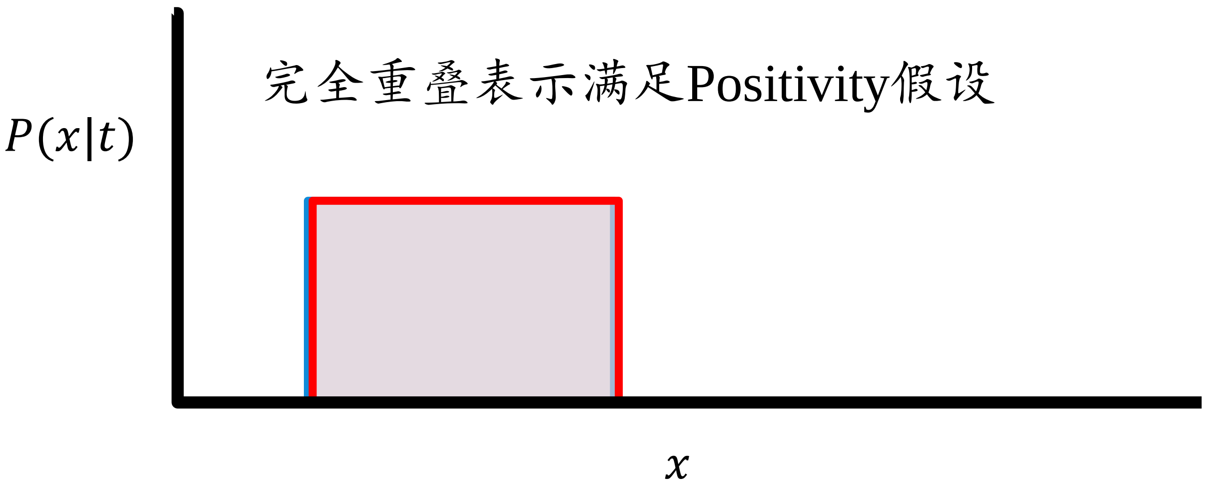
另一种角度：重叠

没有重叠意味着严重违背了Positivity假设



因果推断的基本问题

另一种角度：重叠



因果推断的基本问题

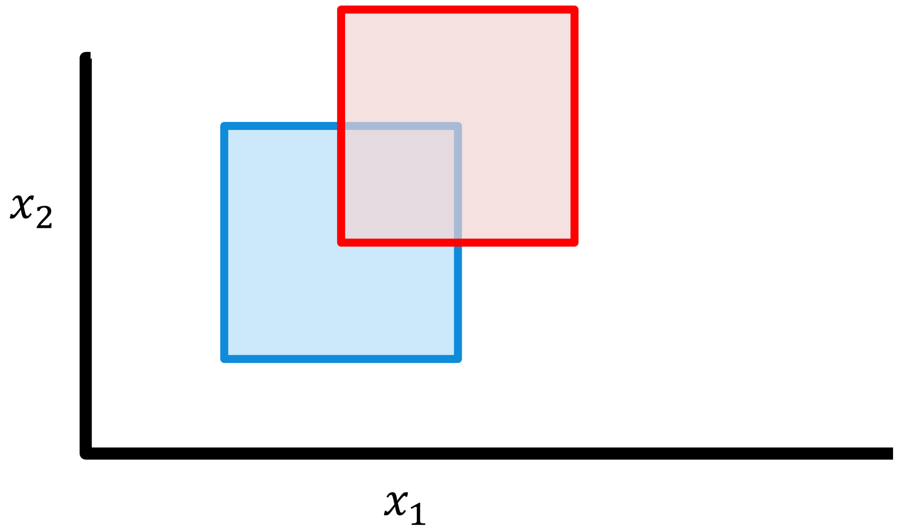
Positivity与Unconfoundedness之间的权衡

50%重叠

一维情况：

25%重叠

二维情况：

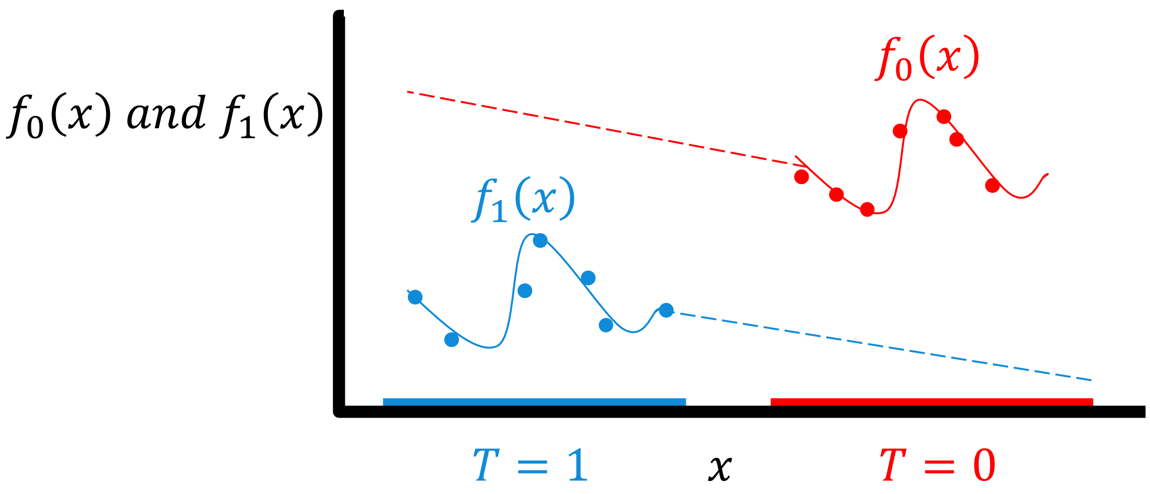


因果推断的基本问题

外推(Extrapolation)

$$\sum_x \mathbb{E}[Y|T=1,X] - \mathbb{E}[Y|T=0,X]$$

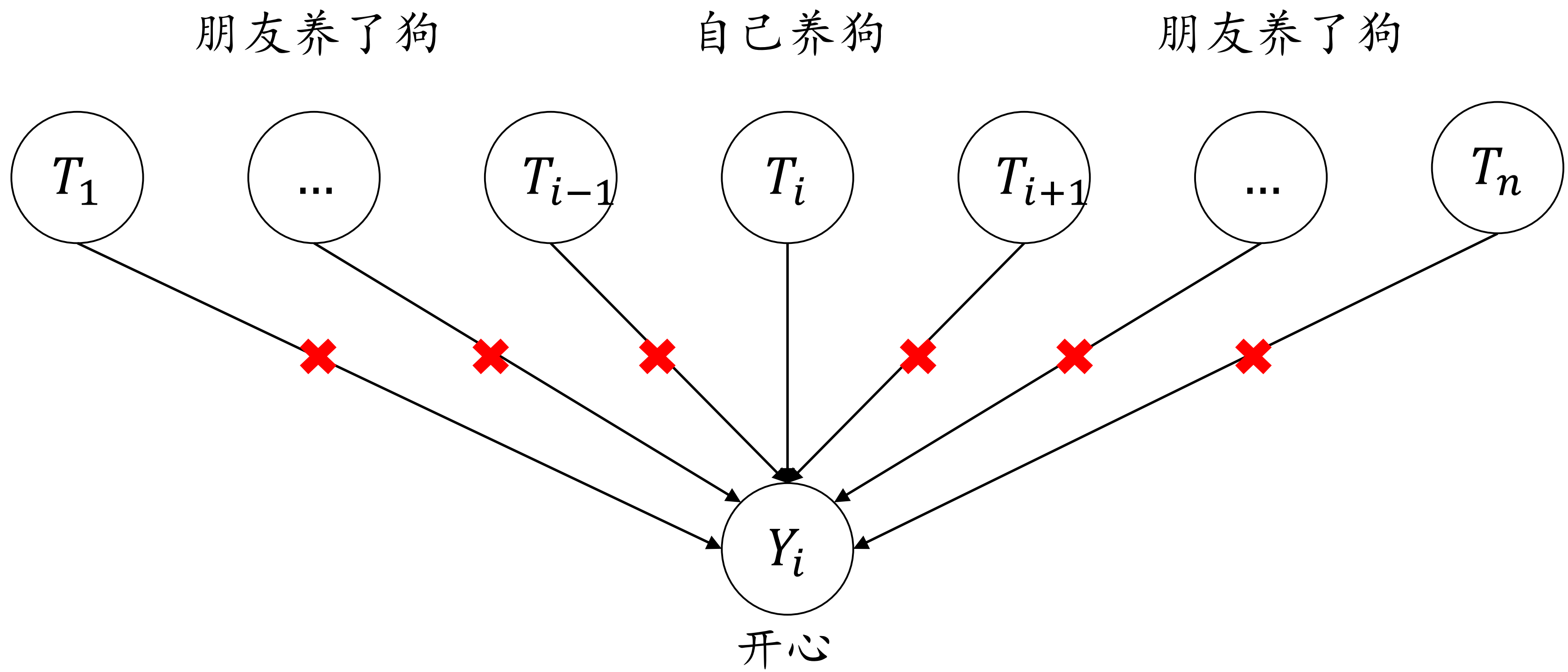
使用$f_1(x)$拟合　使用$f_0(x)$拟合



因果推断的基本问题

无干扰(No interference)

$$Y_i(t_1, \dots, t_{i-1}, t_i, t_{i+1}, \dots, t_n) = Y_i(t_i)$$



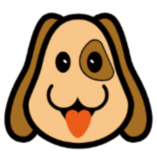
因果推断的基本问题

一致性(Consistency): $T = t \Rightarrow Y = Y(t)$

$T = 1$	$T = 0$
拥有小狗	没有小狗

一只聪明狗

$T = 1 \Rightarrow Y = 1$ 开心



一只傻狗

$T = 1 \Rightarrow Y = 0$ 不开心

违背了一致性假设

因果推断的基本问题

尝试回答以下问题：

1. 要估计平均处理效应，需要满足的四个假设是什么？
2. 为什么违背Positivity假设时需要外推(Extrapolation)？
3. 什么是可识别性(Identifiability)？

因果推断的基本问题

复习一下：

No interference

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$$

linearity of expectation

$$= \mathbb{E}_X[\mathbb{E}[Y(1)|X] - \mathbb{E}[Y(0)|X]]$$

law of iterated expectations

$$= \mathbb{E}_X[\mathbb{E}[Y(1)|T = 1, X] - \mathbb{E}[Y(0)|T = 0, X]]$$

unconfoundedness and positivity

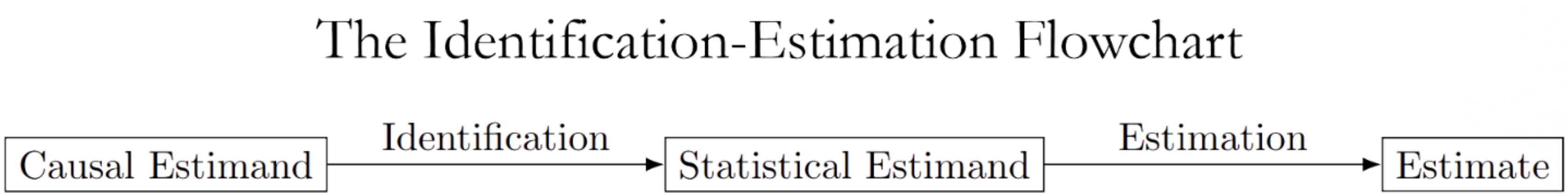
$$= \mathbb{E}_X[\mathbb{E}[Y|T = 1, X] - \mathbb{E}[Y|T = 0, X]]$$

consistency

估计因果效应

Estimands, Estimates, Identification-Estimation Flowchart

- Estimand: 我们想要估计的量
  - Causal estimand(如$\mathbb{E}[Y(1) - Y(0)]$)
  - Statistical estimand(如$\mathbb{E}_X[\mathbb{E}[Y|T=1,X] - \mathbb{E}[Y|T=0,X]]$)
- Estimate: 使用已有数据对Estimand进行估计得到的近似值
- Estimation: 利用数据进行估计的过程



估计因果效应

问题：钠摄入量对血压的影响

研究动机：快节奏的生活使得如今越来越多的人患上了高血压，血管壁长期承受着高于正常的压力会导致冠心病、脑卒中等严重疾病。

数据：

- 案例取自于Luque-Fernandez et al. (2018)
- Outcome Y: 血压（连续型变量）
- Treatment T: 钠摄入量（大于3.5mg为1，否则为0）
- Covariates X: 年龄以及尿液中蛋白质含量
- Simulation: 我们知道“真实”ATE是1.05

估计因果效应

Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X[\mathbb{E}[Y|T = 1, X] - \mathbb{E}[Y|T = 0, X]]$

Estimation: $\frac{1}{n}\sum_x[\underbrace{\mathbb{E}[Y|T = 1, X]} - \underbrace{\mathbb{E}[Y|T = 0, X]}]$

线性回归模型

Estimate: 0.85

$$\frac{|0.85 - 1.05|}{1.05} \times 100\% = 19\%$$

Naive: $\mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0]$

Estimate: 5.33

$$\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$$

估计因果效应

利用线性回归模型的系数

假设Y与T、X满足线性关系: $Y = \alpha T + \beta X$

使用线性回归模型进行拟合: $Y = \hat{\alpha} T + \hat{\beta} X$ $\hat{\alpha} = 0.85$

当T是离散型变量时: $\mathbb{E}[Y(1) - Y(0)]$

当T是连续型变量时: $\mathbb{E}[Y(t)]$ $\longrightarrow$ $\hat{\alpha} = 0.85$

限制：如此估计得到的的因果效应对每个个体来说都是一样的

$Y_i(t) = \alpha t + \beta x_i$

$$Y_i(1) - Y_i(0) = \alpha \cdot 1 + \beta x_i \\ -\alpha \cdot 0 + \beta x_i = \alpha$$

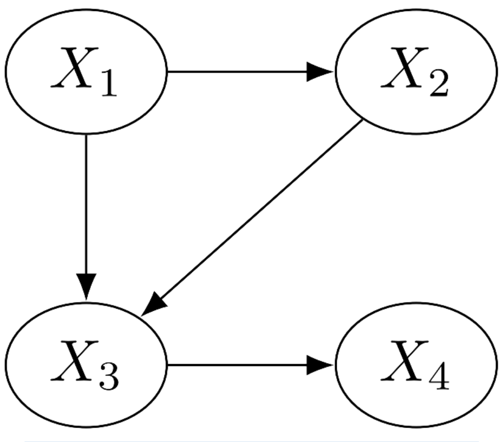
参考Morgan & Winship (2014) 的6.2与6.3节

贝叶斯网络与因果图

简单的对联合分布建模

概率模型（无因果关系）：$P(x_1, x_2, \dots, x_n) = P(x_1)\prod_i P(x_i | x_{i-1}, \dots, x_1)$

$P(x_1, x_2, x_3, x_4) = P(x_1)P(x_2|x_1)P(x_3|x_2, x_1)P(x_4|x_3, x_2, x_1)$



x_1	x_2	x_3	$P(x_4 \mid x_3, x_2, x_1)$
0	0	0	α_1
0	0	1	α_2
0	1	0	α_3
0	1	1	α_4
1	0	0	α_5
1	0	1	α_6
1	1	0	α_7
1	1	1	α_8

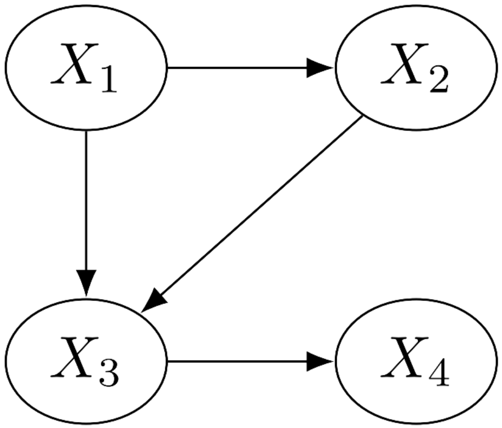
2^{n-1}个参数！

贝叶斯网络与因果图

局部马尔可夫假设(Local Markov assumption)

当在有向无环图中给定了一个节点的所有父节点，那么这个节点就和除了其后代的所有其它节点独立。

$$P(x_1, x_2, x_3, x_4) = P(x_1)P(x_2|x_1)P(x_3|x_2, x_1)P(x_4|x_3, x_2, x_1)$$
$$= P(x_1)P(x_2|x_1)P(x_3|x_2, x_1)P(x_4|x_3)$$

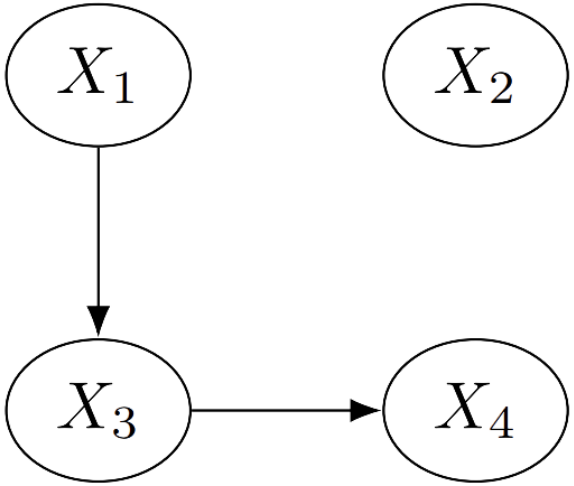


局部马尔可夫假设(Local Markov assumption)

当在有向无环图中给定了一个节点的所有父节点，那么这个节点就和除了其后代的所有其它节点独立。

$$P(x_1, x_2, x_3, x_4) = P(x_1)P(x_2|x_1)P(x_3|x_2, x_1)P(x_4|x_3, x_2, x_1)$$
$$= P(x_1)P(x_2|x_1)P(x_3|x_2, x_1)P(x_4|x_3)$$

现在该如何分解?



$P(x_1)P(x_2)P(x_3|x_1)P(x_4|x_3)$

贝叶斯网络与因果图

贝叶斯网络分解(Bayesian network factorization)

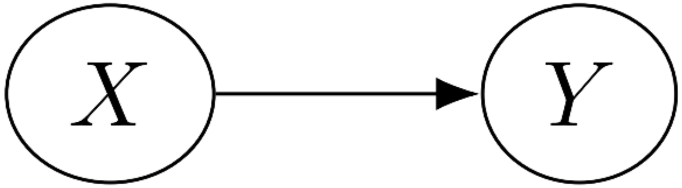
$$P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i | pa_i)$$

Local Markov assumption $\Longrightarrow$ Bayesian network factorization

Local Markov assumption $\Longleftarrow$ Bayesian network factorization

Minimality assumption

1. 当在有向无环图中给定了一个节点的所有父节点，那么这个节点就和除了其后代的所有其它节点独立（局部马尔可夫假设）。
2. 有向无环图中的相邻节点之间相互依赖。

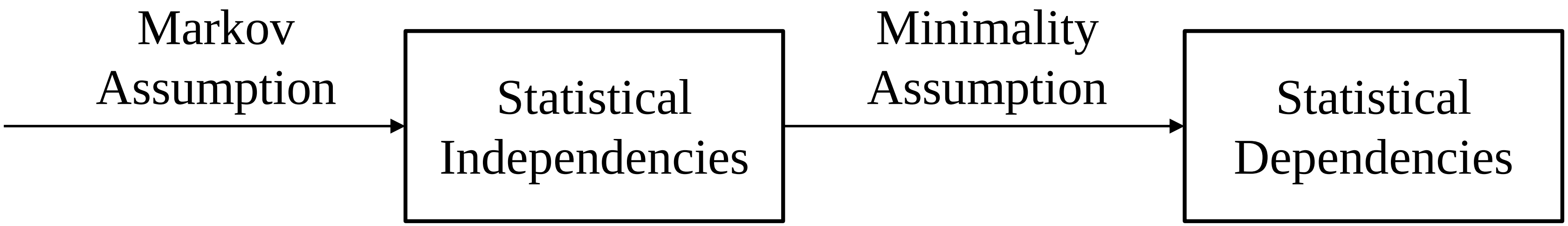


局部马尔可夫假设只告诉了我们变量之间的独立性，仅依据上图和局部马尔可夫假设，我们既可以认为$P(x,y) = P(x)P(y|x)$，也可以认为$P(x,y) = P(x)P(y)$。因此需要引入第二点，保证图中的边对应了变量之间的依赖关系。

Minimality也就是说图中没有多余的无用边，保证有向无环图能够明确表示变量之间的关系。

贝叶斯网络与因果图

假设之间的关联



贝叶斯网络与因果图

尝试回答以下问题：

1. 局部马尔可夫假设和贝叶斯网络分解是什么关系？
2. Minimality假设的两点分别是什么？第二点告诉我们什么？

贝叶斯网络与因果图

什么是cause？

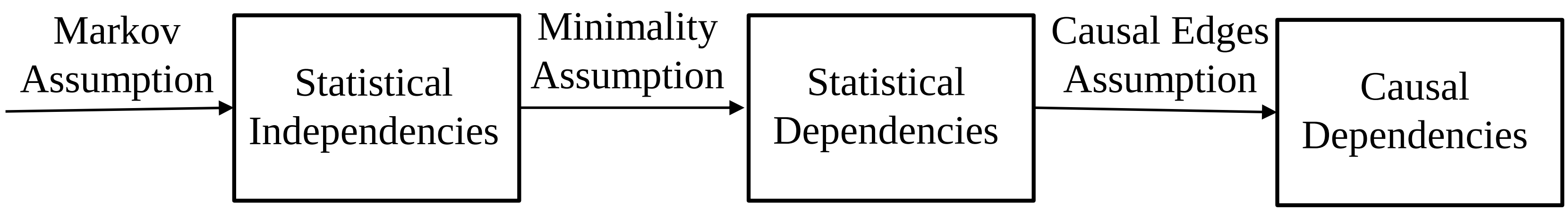
如果变量X的改变会引起变量Y变化，那我们称变量X是变量Y的一个原因(cause)。

Causal edges assumption

在一个有向图中，每一个父节点都是其所有子节点的原因(cause)。

贝叶斯网络与因果图

假设之间的关联



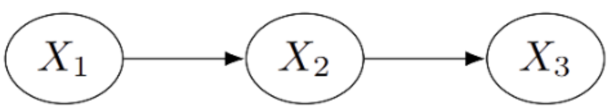
图的基本结构

图的基本组成结构

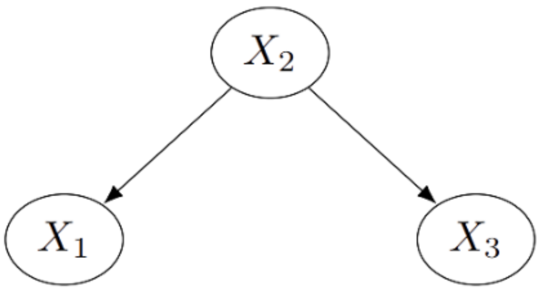
两个节点：

or

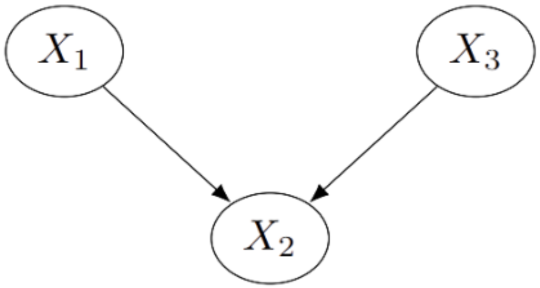
Chain



Fork

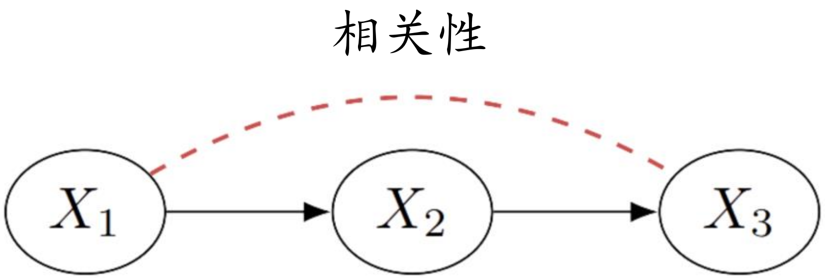


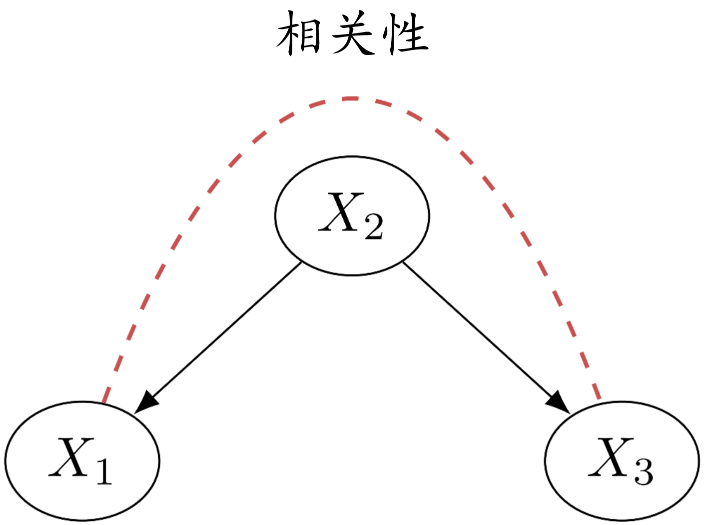
Immorality



图的基本结构

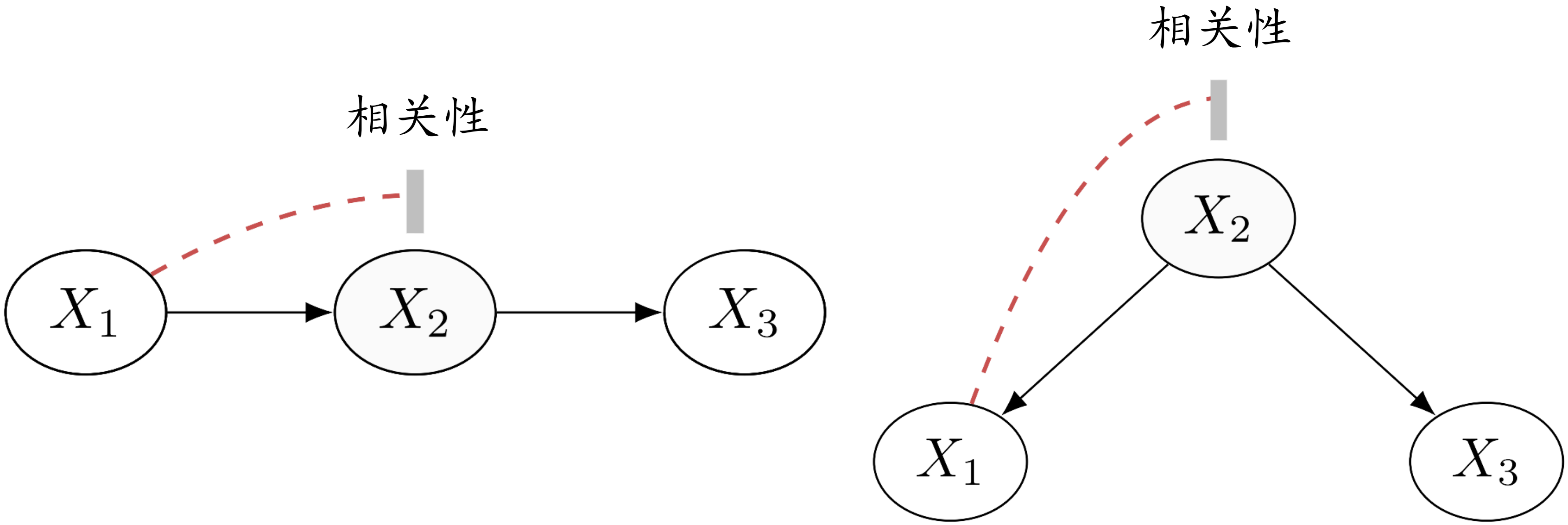
Chains与Forks





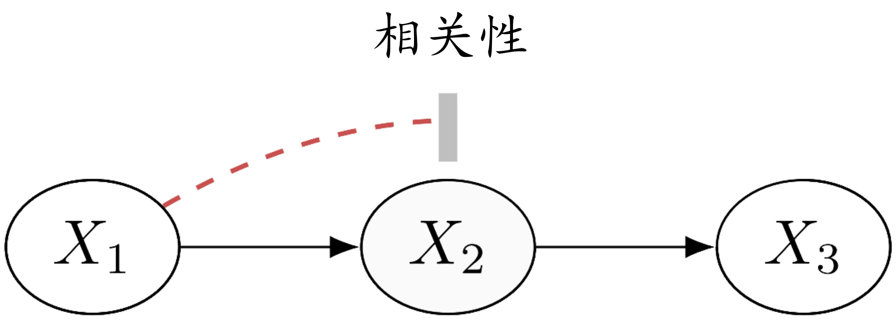
图的基本结构

Chains与Forks

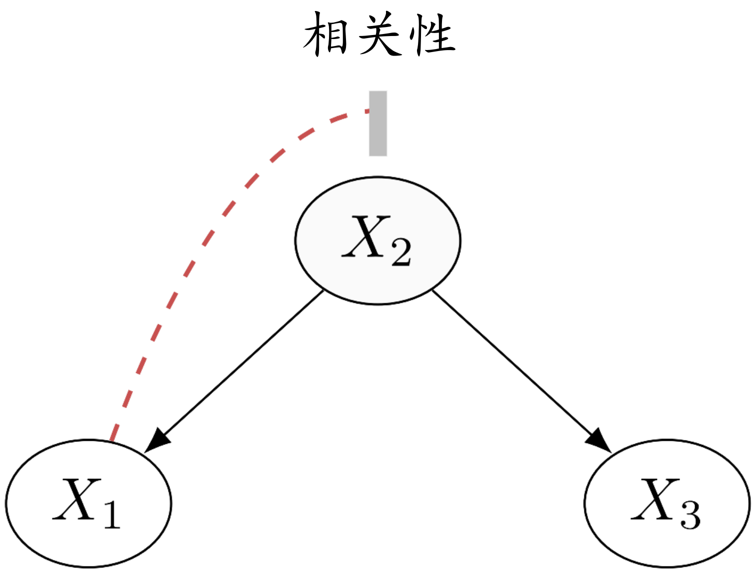


图的基本结构

Chains与Forks



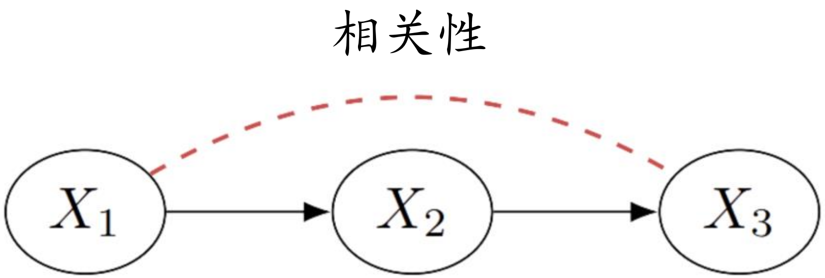
blocked path



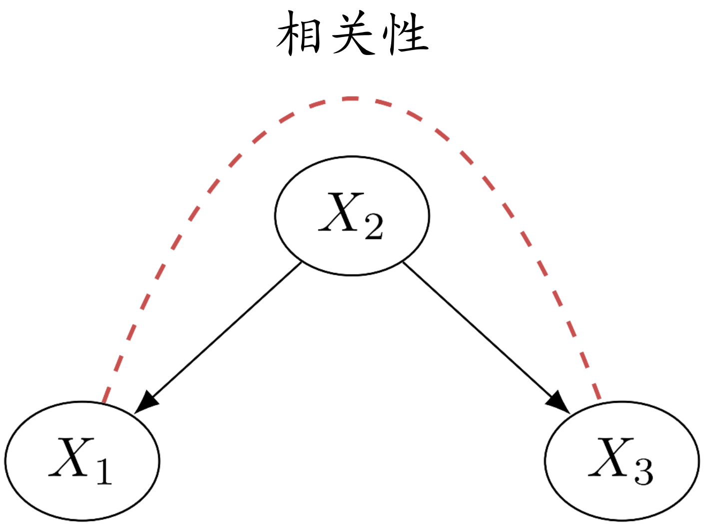
blocked path

图的基本结构

Chains与Forks



unblocked path

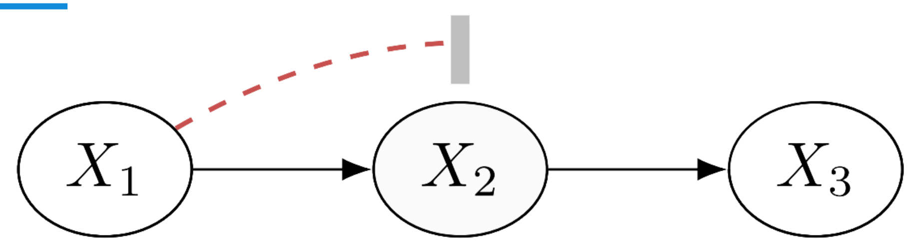


unblocked path

图的基本结构

Chain中条件独立的证明

目标：$P(x_1,x_3|x_2) = \boxed{P(x_1|x_2)}\underline{P(x_3|x_2)}$



1. 贝叶斯网络分解： $P(x_1,x_2,x_3) = P(x_1)P(x_2|x_1)P(x_3|x_2)$

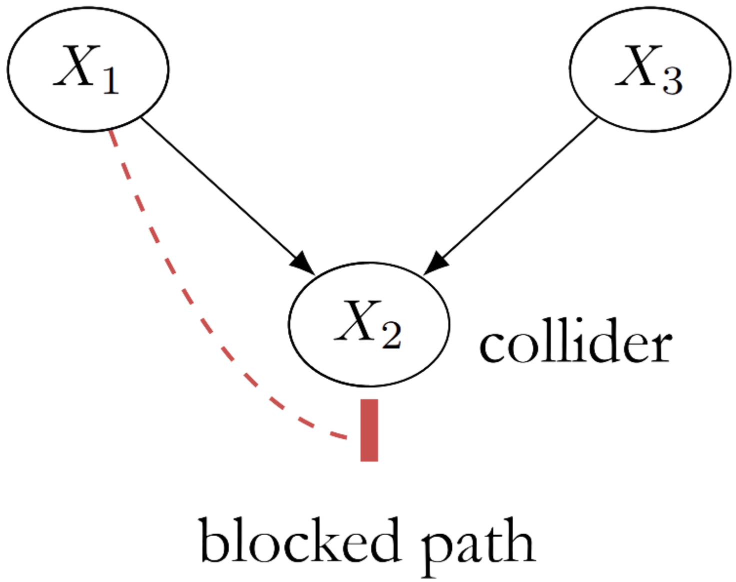
2. 应用贝叶斯公式：$P(x_1,x_3|x_2) = \frac{P(x_1)P(x_2|x_1)\underline{P(x_3|x_2)}}{P(x_2)}$

3. 再次应用贝叶斯公式：$P(x_1,x_3|x_2) = \frac{P(x_1,x_1)P(x_3|x_2)}{P(x_2)}$

$= P(x_1\,|x_2)P(x_3|x_2)$

图的基本结构

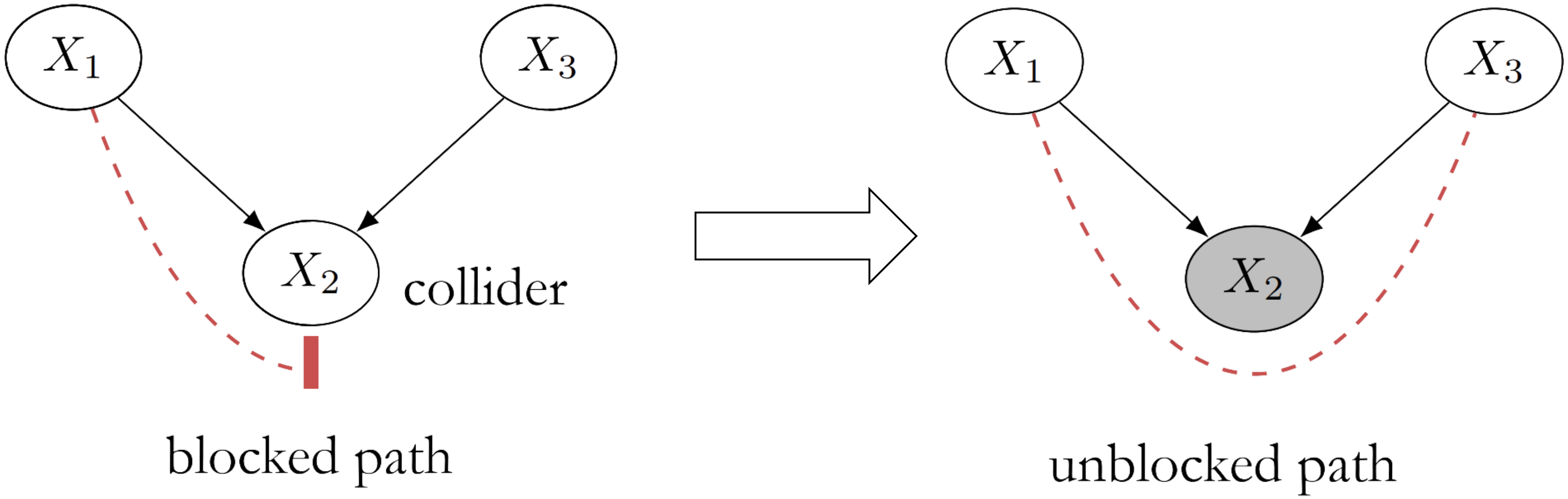
Immoralities



$$
\begin{aligned}
P(x_1, x_3) &= \sum_{x_2} P(x_1, x_2, x_3) \\
&= \sum_{x_2} P(x_1)P(x_3)P(x_2|x_1, x_3) \\
&= P(x_1)P(x_3) \sum_{x_2} P(x_2|x_1, x_3) \\
&= P(x_1)P(x_3)
\end{aligned}
$$

图的基本结构

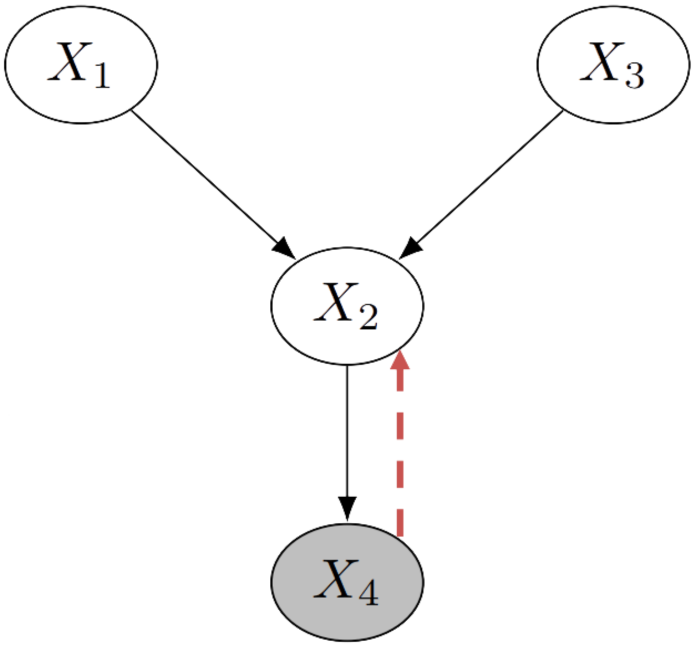
Immoralities: conditioning on the collider



例如：$X_2 = X_1 + X_3$

Conditioning on descendants of colliders

$$X_1 \not\perp\!\!\!\perp X_3 \mid X_4$$



相关与因果关系的流动

阻塞路径(Blocked path)

对于节点X与节点Y之间的一条路径，如果下面条件中任意一条成立，则称X与Y被条件集Z阻塞:

1. 如果W在条件集Z中($W \in Z$)，且这条路径中存在chain ($\ldots \rightarrow W \rightarrow \ldots$) 或者 fork ($\ldots \leftarrow W \rightarrow \ldots$) 结构。
2. 路径中存在collider W，并且W及其后代都不在条件集Z中($W \notin Z, de(W) \not\subseteq Z$)。

D分离(d-separation)

如果两个节点（集合）X与Y之间的所有路径都被条件集Z阻塞，则称X与Y被Z **D分离**。

相关与因果关系的流动

D分离(d-separation)