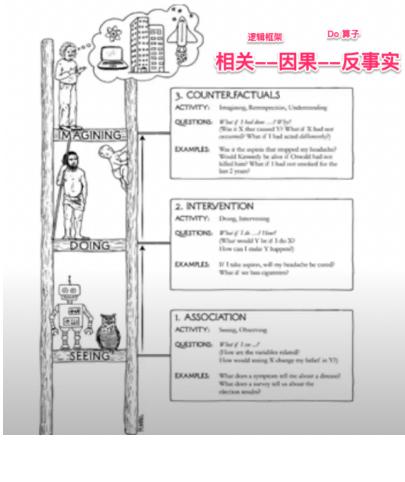
- 因果关系的三层阶梯 The Ladder of Causation
- 2. 统计与因果推断 Statistics and Causal Inference
- 贝叶斯方法 Bayesian Methods

4. 干扰因素 The Confounders

潜在变量 中间变量

- 5. 两个悖论 Two Paradoxes
- 因果与人工智能 Causation and AI
- 7. 总结与感悟 Summary

数据只能用来验证,所做的因果推断是在当前 的因果框架之中的



第三层:反事实 动作:想像,反思 问题:如果我没有做A会发生什么?为什么? 举例:是阿司匹林减轻了我的头痛吗? 第二层:因果

动作:做,干预

问题:如果我做A会发生什么?怎样? 举例:如果我服用阿司匹林,我的头痛会减轻吗?

第一层:相关

动作:看见

问题:如果我看到A,会看到什么? 举例:一项调查能告知我们有关选举结果的哪些信息?

统计与因果推断 Statistics and Causal Inference

● Francis Galton (1822-1911): 提出回归分析和"相关" (Correlation) 的概

计学是一门研究减少数据方法的学科。"

念。尝试用回归解释因果。

Ronald Fisher (1890-1962): 完善回归分析的理论,现代统计学之父。"统

Karl Pearson (1857-1936): 数学统计之父。认为因果只是相关的一个特例。

统计与因果推断 Statistics and Causal Inference

概念。尝试用回归解释因果。

Ronald Fisher (1890-1962): 完善回归分析的理论,现代统计学之父。 "统计学是一门研究减少数据方法的学科。"

Francis Galton (1822-1911): 提出回归分析和 "相关" (Correlation) 的

- Karl Pearson (1857-1936): 数学统计之父。认为因果只是相关的一个特 例。 Sewall Wright (1889-1988): 美国基因学家。在对豚鼠的研究中提出路
- 径分析方法。"无法在没有因果假设的前提下得出因果结论。"

因果图: 基本结构:

贝叶斯方法 Bayesian Method – 贝叶斯网络和因果图

链式:

叉式:

倒叉式:

Sprinkler Rain T 0,1 F 0,5 WetGrass P(W=F) R P(W=T)

Gene

Lung

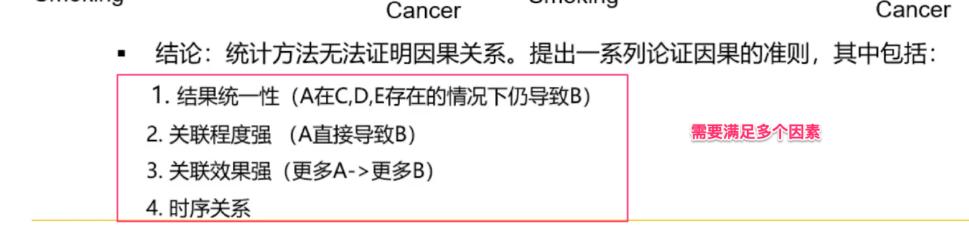
Cloudy

干扰因素 The Confounders

Gene

Lung **Smoking** Smoking

二十世纪中美国社会辩论: 吸烟有害健康吗?



改变选择的结果

Your door

Overall

 $\frac{a_1}{b_1} > \frac{c_1}{d_1}, \frac{a_2}{b_2} > \frac{c_2}{d_2} \implies \frac{a_1 + a_2}{b_1 + b_2} > \frac{c_1 + c_2}{d_1 + d_2}$

87% (234/270)

93% (81/87)

83% (289/350) 87% (234/270) 69% (55/80)

78% (273/350)

83% (289/350)

不改变选择的结果

山羊 山羊

BITTIGER

1号门

跑车

Your door

跑车 山羊 赢 山羊 跑车 赢 表1. 选择1号门时的情境

3号门

山羊

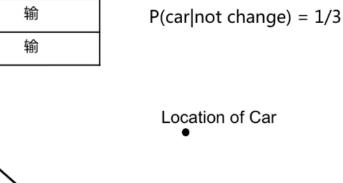
Location of Car

两个悖论 Two Paradoxes

蒙特霍尔问题 (Monty Hall Problem)

2号门

山羊



Patients with

small stones

93% (81/87)

87% (234/270)

P(car|change) = 2/3

Patients with

73% (192/263)

outcome

confounder (blood pressure

outcome

large stones

69% (55/80)



Percutaneous nephrolithotomy 原因:

Treatment A:

Open surgery

Treatment B:

 $\frac{81}{87} > \frac{234}{270}, \frac{192}{263} > \frac{55}{80} \implies \frac{81+192}{87+263} > \frac{234+55}{270+80}$

幸论 Two Paradoxes – 辛普森悖论						
肾结石疗法(研究疗法treatment和疗效outcome的关系)				confounder (stone size)		
	Overall	Patients with small stones	Patients with large stones			
ent A: egery	78% (273/350)	93% (81/87)	73 % (192/263)			

High Blood Pressure Low Blood Pressure

69% (55/80)

73% (192/263)

78% (273/350)

Overall

83% (289/350)

心脏病疗法(研究疗法treatment和疗效outcome的关系)

因果与人工智能	
以深度学习为代表的人工智能学派仍处于因果阶梯第一层。该方法在解决问题过程中缺少透明性和对问题的透彻理解	3. COUNTERFACTUALS ACTIVITY: Irrugining, Remorperate QUESTIONS: What if I had date I i (What if X that date I i vocureed? What if I had BXAMRES: Win or the appent that i Wind Michael had What of I had late 2 years?
强人工智能需要理解因果关系。以作者为代表的贝叶斯学派致力于解决阶梯中第二三层的问题。	2. INTERVENTION ACTIVITY: Doing, Intervening QUESTIONS: What if I do! How' (What would Y be of I to I how' (What would Y be of I to I to we can I make Y hop DXAMPLES: If I take superion, will on What if we has eigenen

1. 解释世界的因果模型 2. 解释自身系统的因果模型 3. 对行为转换为事件的记忆能力

Treatment A: Open surgery Treatment B:

Percutaneous nephrolithotomy

Treatment A: Take Drug

Treatment B: No Drug

总结感悟

很多看似明显的因果关系并非显然

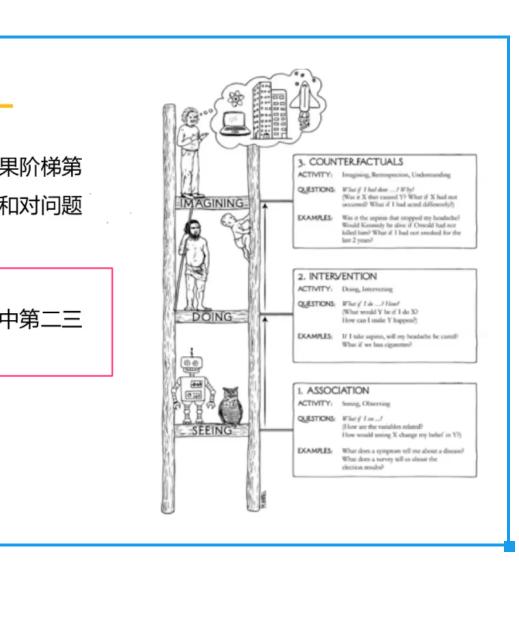
人工智能系统需具备的特点:

认识到统计方法的局限性:

无法通过数据得到因果结论,只能通过数据辅助因果假设

--是否存在潜在变量?--是否存在间接变量?

"比你的数据更聪明"



treatment

treatment