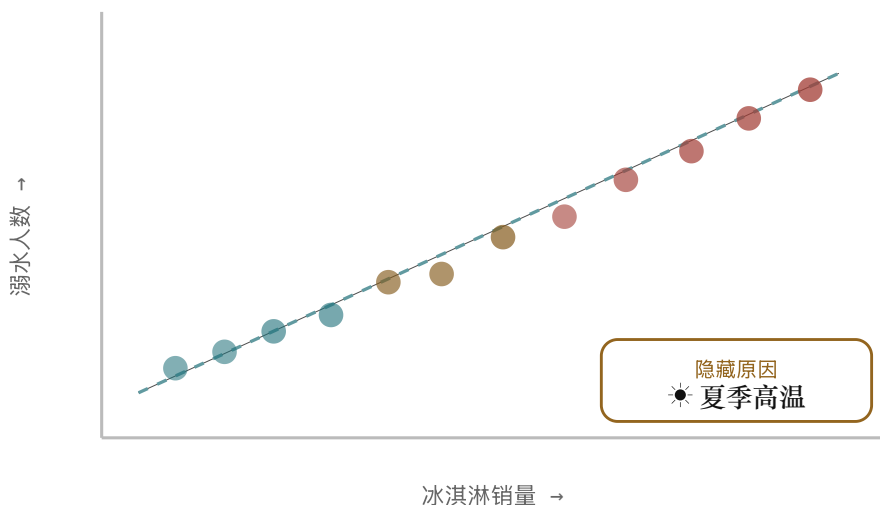


因果性

冰淇淋销量与溺水人数同步起伏。但没有人是因为吃圣代而溺水的。那么，「规律」与「因果」之间的区别究竟在何处？



散点同步上行——这是典型的相关性。它们的颜色代表了图上没有显示的变量：季节。高温既催生了甜筒销量，又让更多人下水消暑。

每年夏天，都有两个数字在同步攀升。随着冰淇淋销量的上扬，溺水人数也随之增加；而当秋凉萧瑟、甜筒不再畅销时，溺水事件也随之平息。将二者绘制在图表上，你会得到一条清晰且笃定的上升曲线——这种典型的相关性足以让粗心的分析师抛出惊悚的标题：「取缔冰淇淋，拯救生命」。然而你早已在直觉中洞悉：这纯属荒唐。没有任何一个人，是仅仅因为吃了一份额圣代就导致了溺水。

你这种「骨子里的直觉」，正是科学中最难诉诸笔墨的事物之一。在这场戏的幕后，潜伏着第三个角色——夏季——它正同时牵动着两根丝线。酷热驱使人们排队购买冰淇淋，同时也驱使人们跳入湖泊与海洋，于是悲剧在其中偶发。这两个可见的数字之所以共舞，仅仅是因为一个不可见的指挥家在幕后调度。今日的主题，便是如何捕捉这位隐形的指挥家，以及现代科学的一项重大发现：**因果性**（Causation）不仅仅是「加强版」的相关性。它位于「因果之梯」的更高处，而那是一个单凭盯着数据看永远无法攀抵的高度。

历经四日，我们的认知工具箱开始严密咬合。**第1日**警示我们留意那些仅凭运气偶然命中的真信念——如停摆的钟。而**虚假相关**（Spurious Correlation）¹正是群体尺度上的「坏钟」：数据虽然「对」得上，原因却错得离谱。**第2日**引入了休谟及其归纳问题；今日我们再次邂逅这位思想家，因为他对因果性的解构与对归纳的抨击实为同一柄利刃的两面。**第3日**理顺了推理的三种模式；而因果发现，正是「带有锋芒的溯因」：它是对最佳因果解释的推断。最后，**第4日**为我们奠定了「干预」与「观察」之别的概率底座：今日的核心结论 $P(y | do(x)) \neq P(y | x)$ ，是本模块迄今为止最重要的不等式。

—— 最古老的问题

休谟踢翻了支撑

从麻烦开始的地方谈起。在1739年的《人性论》中，大卫·休谟提出了一个看似简单得离谱的问题：当一个台球撞击另一个台球，第二个球滚走时，你「究竟」看到了什么？你看到第一个球在移动。你看到它们相互接触。你看到第二个球开始移动。无论你多么努力地观察，你永远无法看到的，是因果本身：即连接这两个事件的「必然联系」、隐藏的力，或是那个代表「因为」的小箭头。

休谟主张，我们实际观察到的只有**恒常结合**（Constant Conjunction）²：一类事件总是稳定地紧随另一类事件之后出现。再加上原因在先（先后性）以及两个事件在时空上接触（接近性），你便拥有了经验所能提供的一切。那种「必然性」的感觉——即第二个球「不得不」移动的感觉——根本不存在于外部世

界。它是一种思维习惯，是重复建立起来的惯常预期，然后像投影一样被投回现实世界。休谟发现，「原因」的两个定义在我们的脑海中纠缠在一起：一个是关于世界的（恒常结合），一个是关于我们的（思维从一个跳跃到另一个的熟练过程）。

这听起来应该很耳熟。这正是[第2日提到的归纳问题](#)换了件马甲。如果因果性仅仅是「此事件总在彼事件之后发生」，那么声称下一次碰撞会与上一次表现一致，就正是对「自然统一性」的一场无法证明的豪赌，而休谟已经证明，这种赌法永远也无法在不陷入循环论证的前提下得到辩护。因果与归纳是同一个伤口。在接下来的两个世纪里，哲学基本上就是在反复抓挠这个伤口。

为什么叫「宇宙的粘合剂」？

这个短语常与休谟联系在一起，但最好将其视为麦基的标题和意象：J.L. 麦基 1974 年的书名为《宇宙的粘合剂》（The Cement of the Universe）。休谟所乐见的讽刺依然存在：这种「粘合剂」是我们永远无法看到的。我们推断出胶水的存在，仅仅是因为砖块总能粘在一起。

—— 四种修正

阐明「更多」的含义

如果因果性不仅仅是恒常结合，那么最显而易见的方法就是说明这「更多」的部分是什么。20 世纪产生了几个严肃的答案——以不同的方式来完成「C 导致 E 意味着……」这个句子。它们与其说是简单的手，不如说是现代因果模型不断借用的透镜。

LEWIS · 1973

反事实 (Counterfactual)³。「C 导致 E」意味着：「若 C 未发生，E 就不会发生。」我们可以借用「可能世界」来理解：想象一个与现实最接近、但 C 缺席的世界，观察 E 是否仍然出现。该定义简洁直观，但需面对「预截」

REICHENBACH · SUPPES ·
CARTWRIGHT

概率性。原因会提高结果发生的概率。莱兴巴赫提出了**共同原因原理** (Common-Cause Principle)⁴：若 A 与 B 相关但互不为因，则必有一个共同原因 C 将其「屏蔽」(Screen Off) ——

(Preemption) 和「过度决定」(Overdetermination) 等复杂因果场景的挑战。

WOODWARD · 2003

干预主义。其核心定义是：若「扰动 C——且仅扰动 C——会导致 E 发生改变」，则 C 是 E 的原因。这种干预无需人类参与：一次干预 (Intervention)⁵ 就像一次外科手术式的精准拨动，因此即便无人按下按钮，火山喷发依然是火山灰产生的原因。这是珀尔理论体系在哲学上的孪生兄弟：下文中，珀尔会把这种操作形式化为 do(C) 算子和图上的剪箭头操作。

即在固定 C 的条件下，A 与 B 的相关性会消失（如前述夏季的例子）。

核心脉络

一个问题，多个视角。反事实视角关注「若非如此会怎样？」；概率视角关注「在固定干扰项时，它是否增加了胜算？」；干预视角关注「拨动它会引起什么变化？」。珀尔框架为这些直觉提供了一套统一的数学语言。

注意卡特赖特 (1979) 是如何完善概率论述的，因为她的修正正是今天的转折点。原因确实提高了结果的概率——但仅限于在「因果同质」的背景下，即所有其他原因都保持固定的情况下。忘记这个前提，你就会直接掉进统计学中最精妙的陷阱。

—— 陷阱

辛普森悖论：当数字发生彻底反转

若非亲眼所见，简直难以置信：某种疗法可能对小结石效果更好，对大结石效果也更好，但总体表现却更差。不是「看起来更差」——而是在汇总数据上确实更差。下方这组真实的肾结石数据会展示这种反转是如何发生的。

其机制总是一个「潜伏变量」在你所比较的组别之间分布不均。在肾结石数据 (Charig 等人, 《英国医学杂志》, 1986 年) 中, 外科医生主要将微创的疗法 B 用于「容易处理」的小结石, 而将传统的开放式手术 A 留给「困难」的大结石。因此, B 的整体成功率因其更容易处理的病例而显得虚高。如果按结石大小拆分——正如卡特赖特所要求的, 通过分层比较固定混杂因素 (Confounder)⁶——A 在两组中都会胜出。

下面的「反转机器」会把这件事压成一张图。每一个方块代表一名患者，而且永远不移动。颜色表示所接受的疗法；所在行表示病例是小结石还是大结石；填充表示治疗是否有效。汇总视图比较两列疗法；条件化视图保留同一张图，只是在每一行内部先比较 A 与 B。这就是条件化的基础：先让比较对象可比，再谈因果判断。

肾结石辛普森反转

视图	疗法 A	疗法 B	表观赢家
所有患者	273 / 350 = 78.0%	289 / 350 = 82.6%	B
小结石	81 / 87 = 93.1%	234 / 270 = 86.7%	A
大结石	192 / 263 = 73.0%	55 / 80 = 68.8%	A

汇总比较之所以反转，是因为 B 被更多地用于较容易处理的小结石病例。

教训是惨痛的：你无法从数字表格中读出因果关系。完全相同的数字，根据一个甚至可能不在电子表格中的变量，可以支持截然相反的结论。这引出了一个统领后续所有内容的问题——如果光靠数据无法告诉你答案，那么「什么」能告诉你？答案来自一位计算机科学家，他在 20 世纪 80 年代致力于让机器在不确定性中推理：你需要添加数据中不包含的东西——一个关于「箭头指向何处」的模型。

—— 因果革命

珀尔的梯子，以及改变一切的动词

朱迪亚·珀尔（Judea Pearl）获得了 2011 年的图灵奖——通常被称为计算机界的诺贝尔奖——「因其通过开发概率和因果推理演算，对人工智能做出的基础性贡献」。他的核心意象在他 2018 年的科普著作《为什么》（The Book of Why）中得以普及，即**因果之梯**（Ladder of Causation）⁷。它分为三层，每一层都比下一层要求更深刻的问题，也要求更有力的证据。珀尔喜欢讽刺道，大多数统计学和几乎所有的机器学习从未离开过最底层。

亲自攀爬：

珀尔的因果之梯

层级	动词	记号	问题
1. 关联	观察	$P(Y X)$	在观察到 X 的案例中，结果有何不同？
2. 干预	干预	$P(Y do(X))$	如果通过干预设定了 X，会发生什么？
3. 反事实	想象	$P(Y_x X')$	在已知现实结果的前提下，若过去的操作有所不同，结果会如何变化？

在最后一行的记号中， Y_x 代表「在 X 被干预设定为 x 的平行世界里，变量 Y 的潜在值」；而 X' 与 Y' 则代表我们已经观测到的现实事实。

do-算子：观察不等于干预

这是整个现代因果领域的概念枢纽，值得仔细体味。对于变量 X ，你可以进行两种截然不同的操作。

你可以对它进行条件化 (Conditioning) ⁸——记作 $P(Y \mid X = x)$ 。这代表：在所有 X 「恰好」等于 x 的既定案例中， Y 的分布如何？你只是在筛选现有的数据池。这就是「观察」 (Seeing)。

或者你可以进行干预——在珀尔的记号中写为 $P(Y \mid \text{do}(X = x))$ 。这意味着：介入其中，强迫每个人的 X 都等于 x ，将 X 与其通常的原因切断，然后观察 Y 。这就是「干预」 (Doing) ——这正是随机实验所做的事情。

当存在混杂因素时，这两个数值就会分离，而它们之间的差距就是「偏差」。在那些「恰好」买了大量冰淇淋的人中，溺水确实更常见（因为他们是夏天的那批人）——所以「观察」到的数值很高。但是，如果你强迫一组随机样本购买冰淇淋（无论天气如何），溺水人数并不会变化——「干预」后的数值是平整的。下方的机器让你能够调节混杂因素的强度，观察这两个数值如何分道扬镳。

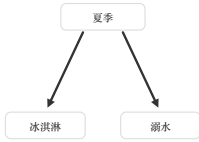
观察不等于干预

在这个模型世界中，观察到的关联等于真实直接效应加上来自夏季的偏差项。在默认设置下，真实直接效应为 +0.05，但观察到的关联要大得多，因为夏季同时推高了冰淇淋购买量和游泳溺水人数。

量	问题含义	默认读数
$P(\text{溺水} \text{购买}) - P(\text{溺水} \text{未购买})$	观察：比较现有的购买者与非购买者。	真实效应加上混杂偏差。
$P(\text{溺水} \text{do}(\text{购买})) - P(\text{溺水} \text{do}(\text{未购买}))$	干预：直接设定购买行为并切断其通常的原因。	仅包含真实直接效应。

右侧图中被切断的箭头就是 do-算子的直观表现。干预不仅仅是「看」X——它删除了指向 X 的箭头，并用你的「手」取而代之。「夏季 → 冰淇淋」的链接被切断，因此夏季再也无法利用冰淇淋作为「后门路径」来伪造对溺水的效应。留下来的只有真实的因果关系。

珀尔为我们提供了一套在纸上进行此类操作的语法。一个结构因果模型（Structural Causal Model）⁹ 将变量绘制为有向无环图（Directed Acyclic Graph，简称 DAG）¹⁰——即由方框与箭头组成且无循环的图——其中三种命名模式承担了大部分工作。一个叉形结构（Fork， $X \leftarrow Z \rightarrow Y$ ）是一个混杂因素；通过对 Z 取条件来阻断它。一个链形结构（Chain， $X \rightarrow Z \rightarrow Y$ ）是一个中介因子。而一个对撞结构（Collider， $X \rightarrow Z \leftarrow Y$ ）则是陷阱：X 和 Y 本身无关，直到你对它们的共同结果 Z 进行条件化处理，此时一个幻影般的相关性便会跃然纸上。这就是为什么「控制你能测量的一切」并非严谨而是鲁莽——对一个对撞因子取条件，你就会「制造」出你原本试图消除的偏差。



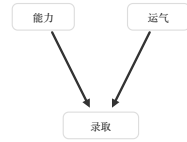
叉形 / 混杂

夏季同时改变暴露与结果。比较前先对夏季取条件。



链形 / 中介

焦油承载吸烟的部分效应。对焦油取条件会改变问题。



对撞结构

能力与运气在录取处相遇。对录取取条件会打开伪相关。

前门准则

珀尔的前门准则（Front-Door Criterion）¹¹是该框架中最巧妙的招式之一：有时即使存在未测量的基因混杂因素，你仍然可以估算吸烟对癌症的效应，只要你能测量中间的一个完整中介因子（例如肺部的焦油沉积）。计算分三步：估计吸烟如何改变焦油；在考虑吸烟的同时估计焦油如何改变癌症；再把这些部分按观察到的吸烟分布进行加权平均。隐藏的基因混杂因素虽然影响吸烟和癌症，但不会进入最终前门表达式里的那两个可测片段。这是一种真正的「从观察数据中获取因果答案」的手段——但前提是你必须提供一张能够说明焦油是完整中介因子的因果图。天下没有免费的午餐：假设只是从电子表格转移到了因果图中。

—— 前沿 · 2026

三个活跃领域——以及前沿校准器

现在来谈谈那个催生了数千篇论文、甚至开创了一个行业的问题：你能仅从观察中推断出因果吗？2026年的坦诚回答是精确的「部分可以——但存在一堵已被证明的墙」。下方的每项声明都标明了其可信度。

前沿 01 [已确立] [已确立]

圈定领域的两个定理

这是今天最稳固的基石——它不是可能被推翻的经验发现，而是「数学证明」。首先，do-演算具有完备性。珀尔的四条重写规则只要在可能的情况下，就能将 `do()` 表达式转化为普通的概率表现形式；Shpitser & Pearl (2006) 以及 Huang & Valorta (2006) 分别证明了，如果这些规则「无法」消除 do-算子，那么没有任何方法可以做到——该效应在当前观察数据加上那张图的前提下确实是不可识别的。珀尔称这终结了「非参数识别的篇章」。对于「观察何时能替代干预」这一问题，这是一个简洁、永久的答案。

DO-演算的重写长什么样

- 删除无关观察：如果 Z 已经阻断了 X 到 Y 的所有开放路径，那么知道 X 不再增加任何信息： $P(Y | X, Z) = P(Y | Z)$ 。
- 把行动换成观察：在经过正确的变量调整后，设定 X 与观察到 X 可以回答同一个问题： $P(Y | do(X), Z) = P(Y | X, Z)$ 。
- 删除无关行动：如果在固定 X 后，对 Z 的干预已经无法影响 Y，就可以将其删除： $P(Y | do(X), do(Z)) = P(Y | do(X))$ 。

这些只是说明性的例子。关键在于这种逻辑：因果图会授权你精确地擦掉或交换哪些符号。

其次，是另一边的墙：马尔可夫等价天花板。仅利用观察数据中的条件独立模式，某些不同的因果图在「数学证明上是不可区分的」。例如 $X \rightarrow Y \rightarrow Z$ 、 $X \leftarrow Y \leftarrow Z$ 和 $X \leftarrow Y \rightarrow Z$ 都暗示了同一个事实（「一旦知道 Y，X 和 Z 就是独立的」），因此再多的此类数据也无法区分它们。它们构成了一个马尔可夫等价类 (Markov Equivalence Class)¹²。只有「对撞结构」($X \rightarrow Y \leftarrow Z$) 表现特殊：在原始数据中 X 与 Z 相互独立，但一旦对它们的共同结果 Y 进行条件化处理，二者之间就会产生伪相关。结论是冷酷的：在无假设的情况下，单靠观察永远无法提供唯一的因果图——只能提供一类候选。这两项结果在最强的意义上都是「已确立」的：它们是定理。

前沿 02 [前景可期] [已确立]

从静态数据中识别因果方向——以及为什么实验依然为王

那么，这堵墙就是终点吗？并非如此——你可以通过引入独立性测试本身不使用的额外假设来翻越它。**LiNGAM** (Shimizu 等, 2006)¹³证明了：在满足线性关系且无隐性混杂的条件下，只要随机噪声服从「非高斯分布」，因果方向就是可识别的——这种非对称性终结了 $X \rightarrow Y$ 与 $Y \rightarrow X$ 之间的「平局」局面。在真实方向上，剩余噪声与原因相互独立；在错误方向上，残差仍会留下可检测的依赖痕迹。**加性噪声模型** (Additive-Noise Models)¹⁴ (Hoyer, Janzing, Mooij, Peters & Schölkopf, NeurIPS 2009) 将这一理论扩展到了非线性因果对。在标准的 Tübingen 因果对 (Tübingen Cause-Effect Pairs)¹⁵ 基准测试中——108 对具有已知真实情况的数据，如「海拔 \rightarrow 温度」——一个强大的系统报告了约 83% 的准确率 (Mosaic, Wu & Fukumizu 2020)，这证明了一些「通常被认为不可能」的事情在部分程度上是可行的。但请注意这种技巧的形式：你只有通过「假设」非高斯性或加性才能翻过墙。这些假设单靠同一份观察分布是无法验证的——而且对于整个基于约束的进路来说，那个起支撑作用的假设——**忠实性** (Faithfulness)¹⁶，在有限样本中可能会悄无声息地失效。[前景可期]

这就是为什么金标准依然简单得残酷：做实验。一次**随机对照试验**

(Randomized Controlled Trial, 简称 RCT)¹⁷通过随机分配 X 在现实中执行 **do(X)**，一举切断了所有后门路径。当你无法进行随机化时，经济学界的**可信度革命** (Credibility Revolution)¹⁸致力于寻找能够模拟随机化过程的「自然实验」——例如工具变量、断点回归或双重差分法。该项目为大卫·卡德、约书亚·安格里斯特和圭多·因本斯赢得了 2021 年诺贝尔经济学奖（我们将在第 152 日详细讨论）。[已确立]

前沿 03 [前景可期] [争议/炒作]

机器能做到吗？因果机器学习与「因果鸚鵡」

最热门也最模糊的前沿。[因果表征学习](#)（Causal Representation Learning）

（Schölkopf 等，2021）¹⁹提出了一个深刻的问题：经典的因果发现假设变量是现成给定的，但现实世界是以像素和词语的形式呈现的。网络能否「学习」到高层因果变量——这是否能让模型对[分布偏移](#)（Distribution Shift）²⁰更具鲁棒性，而这正是当今模型通常欠缺的？这是一个严肃且活跃的项目，其最大的承诺目前仍属于「前景可期」，尚未大规模兑现。

接着是最具争议的问题：大型语言模型能进行因果推理吗？Kıcıman, Ness, Sharma & Tan (2023) 报告称，GPT-4 在 Tübingen 成对因果方向任务上达到了 97% 的准确率——比之前的最佳成绩提高了 13 个百分点——并在反事实评分上表现强劲，认为仅凭记忆无法解释这一点。反驳意见也来得很快：Zečević 等人在《因果鸚鵡》（Causal Parrots, 2023）中针锋相对地指出：大模型只是在「谈论」因果，而非「理解」因果——它们只是在复述训练语料中浸泡过的因果事实，而非真正执行珀尔式的逻辑推理。2024–25 年的综合观点（例如 Jin 等，ICLR 2024，《大模型能从相关性中推断因果吗？》）采取了折中立场：模型是强大的「因果知识检索器」，在第 1 层问题上通常表现出色，但真正的考验是全新的第 2/3 层干预和反事实结构。回想一下[第 1 日的盖梯尔陷阱](#)的新变体：一个输出正确因果声明的 LLM，如果其原因与因果结构无关，那么它是对的，但它「知道」吗？再回想一下[第 3 日](#)：背诵记忆的事实并不是溯因。关于「LLM 在珀尔意义上具有因果推理能力」的结论：「[争议/炒作](#)」它对于辅助因果分析很有用，但尚未确立其作为因果推理器的地位。因果机器学习底层的开源工具（DoWhy, EconML）是真实的，且建立在上述定理之上，但关于「因果 AI 将很快取代普通相关性机器学习」的营销口号仍超前于证据。

—— 开放性问题

尚未定论的问题

- 是否存在一种「正确」的因果理论？或者反事实、概率性和干预主义论述各抓住了不同的侧面，且都无法还原为其他论述？没有任何分析能避开所有的反例（预截、过度决定、情境一致性）。
- 我们在多大程度上可以信任忠实性？即假设现实系统永远不会有精确抵消的因果路径，这很方便但无法测试——而具有反馈和稳态机制的生物学可能会

经常违反它。

- 因果变量能否从原始数据（像素、语言）中学习，而不是手动交给算法？这甚至是一个定义良好的问题吗？因为将世界划分为变量的「正确」方式本身可能就取决于视角。
- 大型模型是否构建了内部因果世界模型，还是仅仅建立了因果言论的统计模型？答案直接指向第 138-145 日，以及预测是否等同于理解的问题。
- 箭头从何而来？这里的每种方法都需要「某种」因果输入——一张图、一个假设或一个实验。休谟的幽灵依然在发问：这些输入究竟是从世界中读取的，还是被我们带入世界的？

◆ 今日三句话总结

核心思想

因果性不是更强的相关性，而是一种「不同类型」的事物，它位于因果之梯的更高层级——「干预」和「想象」高于单纯的「观察」——因此只要存在混杂因素， $P(Y|\text{do}(X)) \neq P(Y|X)$ 就会成立，而在不引入因果假设的情况下，光靠盯着观察数据无法消除这一鸿沟。

最佳类比

冰淇淋和溺水人数同步上升，而夏季在幕后同时牵动着这两根弦；do-算子就像一把剪刀，切断了指向变量的那根弦，让你能看清它真正驱动的是什么。

实时争议

因果能否仅从观察中推断？没有额外假设时，证明只能走到「马尔可夫等价类」；引入假设（如 LiNGAM、加性噪声模型）后可部分恢复方向。此外，2026 年还在激烈争论：LLM 究竟是真正具备因果推理能力，还是只是在模仿因果言论。

今日线索 › 信息（图表是数据所缺乏的额外信息；do-算子量化了观察与干预的区别）· 计算（do-演算作为一种完备的「算法」；因果发现作为一种搜索）· 涌现（因果结构是原始相关性之上的更高层级）——并回溯至 [第 1 日](#)（凭运气正确）、[第 2 日](#)（休谟）、[第 3 日](#)（溯因）以及 [第 4 日](#) ($P(y|x)$)。

明日 → 第 06 日

统计学与不自欺的艺术

今天我们看到了混杂因素如何反转结论。明日我们将遇到更微妙的敌人：你自己。硬币实验中的 p -值操纵、路径分叉花园、效应量与对「显著性」的崇拜——我们刚刚遇到的对撞因子陷阱将作为诚实的分析师欺骗诚实听众的最简单方式之一再次出现。带着今天的本能，每一次都去追问：「那张图表上没有显示的变量是什么？」

说明

1. 虚假相关指两个测量变量间存在真实的统计关联，但这种关联并非源于直接的因果纽带，而是由其他因素诱发的。
2. 恒常结合意味着某种类型的事件反复紧随另一种类型的事件发生。
3. 「反事实」探讨的是如果过去的事实发生了改变，原本的结果会如何变化。
4. 该原理指出，两个变量间异常的相关性通常源于直接因果关系，或是一个能同时解释二者的共同前置原因。
5. 干预是一种理想化的改变，它在强制设定某个变量值的同时，切断了该变量与其通常原因之间的联系。
6. 混杂因素是同时影响表现原因和表现结果的第三个变量，会导致二者产生虚假的同步变动。
7. 珀尔的梯子将关联、干预和反事实推理分为三个层级，每个层级对问题的深度和证据的要求都比下一层更高。
8. 条件化（或称取条件）意味着将分析视角收窄至该变量已具有特定观测值的子样本上。
9. 结构因果模型描述了变量是如何从其他变量与背景噪声中生成的。
10. 有向无环图（DAG）是由箭头组成的网络，且不包含任何有向环。
11. 当直接因果路径被不可测因素混杂时，前门准则可以通过测量中间的中介因子来识别因果效应。
12. 马尔可夫等价类是一组暗示相同条件独立性的因果图，因此仅凭此类观察数据无法在它们之间做出选择。
13. LiNGAM 是「线性非高斯无环模型」的缩写，是一种因果发现方法。
14. 加性噪声模型假设效应等于原因的函数加上独立的噪声，这种不对称性有时可以识别因果方向。
15. Tübingen 基准是一组由专家标记了因果方向的真实世界变量对。
16. 忠实性是指数据中的独立性反映了图结构，而非因果路径之间精确的相互抵消。
17. 随机对照试验通过随机分配参与者到不同条件，从而使背景原因在平均意义上达到平衡。
18. 这场革命将研究重心从复杂的回归模型转向了模拟随机实验的设计，如寻找能使受试者「近似随机」接受处理的自然场景。
19. 因果表征学习试图从原始数据（如像素、文本或传感器流）中学习有用的因果变量。

20. 分布偏移意味着模型部署后看到的数据与其学习时的数据存在差异。

—— 证据

来源与深入阅读

1. Hume, D. (1739–40). *A Treatise of Human Nature*, Book I, Part III; and (1748) *An Enquiry Concerning Human Understanding*, §VII. – 恒常结合；无可见的必然联系。
2. Mackie, J. L. (1974). *The Cement of the Universe: A Study of Causation*. Oxford University Press. – INUS 条件；书名短语。
3. Lewis, D. (1973). "Causation." *Journal of Philosophy* 70(17): 556–567. doi:10.2307/2025310. doi.org/10.2307/2025310 另见 Lewis, *Counterfactuals* (Blackwell, 1973); 修订后的 "influence" 论述 (2000)。
4. Reichenbach, H. (1956). *The Direction of Time*. University of California Press. – 共同原因原理与屏蔽。
5. Suppes, P. (1970). *A Probabilistic Theory of Causality*. North-Holland. – 表面原因与虚假原因。
6. Cartwright, N. (1979). "Causal Laws and Effective Strategies." *Noûs* 13(4): 419–437. – 仅在因果同质背景下的概率提升。概览
7. Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press. – 干预主义/操控主义理论。
8. Charig, C. R., Webb, D. R., Payne, S. R. & Wickham, J. E. A. (1986). "Comparison of treatment of renal calculi by open surgery, percutaneous nephrolithotomy, and extracorporeal shockwave lithotripsy." *British Medical Journal* 292: 879–882. – 肾结石辛普森悖论数据。
9. Simpson, E. H. (1951). "The Interpretation of Interaction in Contingency Tables." *JRSS B* 13: 238–241. Blyth, C. R. (1972), *JASA* 67: 364–366 (提出 "Simpson's paradox" 一词)。Yule, G. U. (1903) 关于虚假相关。
10. Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge University Press. 以及 Pearl, J. & Mackenzie, D. (2018). *The Book of Why*. Basic Books. – 因果之梯；do-演算；后门/前门。

11. Shpitser, I. & Pearl, J. (2006). "Identification of Joint Interventional Distributions in Recursive Semi-Markovian Causal Models." *AAAI*. & Huang, Y. & Valtorta, M. (2006). "Pearl's Calculus of Intervention Is Complete." *UAI*. – do-演算的完备性（已证）。
12. ACM (2012). 2011 A.M. Turing Award – Judea Pearl. amturing.acm.org/award_winners/pearl
13. Spirtes, P., Glymour, C. & Scheines, R. (2000). *Causation, Prediction, and Search*, 2nd ed. MIT Press. – PC 与 FCI 算法；马尔可夫等价。
14. Shimizu, S., Hoyer, P. O., Hyvärinen, A. & Kerminen, A. (2006). "A Linear Non-Gaussian Acyclic Model for Causal Discovery." *JMLR* 7: 2003–2030. 参考代码
15. Hoyer, P., Janzing, D., Mooij, J., Peters, J. & Schölkopf, B. (2009). "Nonlinear causal discovery with additive noise models." *NeurIPS*. Mooij, J. et al. (2016). "Distinguishing Cause from Effect Using Observational Data." *JMLR* 17(32): 1–102. Wu, P. & Fukumizu, K. (2020). "Causal Mosaic: Cause-Effect Inference via Nonlinear ICA and Ensemble Method." *Proceedings of Machine Learning Research* 108: 1157–1167. – 加性噪声模型、Tübingen 成对基准与 Mosaic 结果。
16. Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A. & Bengio, Y. (2021). "Toward Causal Representation Learning." *Proceedings of the IEEE* 109(5): 612–634. doi:10.1109/JPROC.2021.3058954.
17. Kıcıman, E., Ness, R., Sharma, A. & Tan, C. (2023). "Causal Reasoning and Large Language Models." arXiv:2305.00050; TMLR 2024. & Zečević, M., Willig, M., Dhani, D. S. & Kersting, K. (2023). "Causal Parrots: Large Language Models May Talk Causality But Are Not Causal." *TMLR*. arXiv:2308.13067. & Jin, Z. et al. (2024). "Can Large Language Models Infer Causation from Correlation?" *ICLR*.
18. The Royal Swedish Academy of Sciences (2021). The Sveriges Riksbank Prize in Economic Sciences – Card, Angrist & Imbens. nobelprize.org/prizes/economic-sciences/2021
19. Stanford Encyclopedia of Philosophy: "Causation" / "Counterfactual Theories of Causation" / "Causal Models" / "Probabilistic Causation." plato.stanford.edu/entries/causation-counterfactual

第 05 日结束 · 还剩 175 次深入