

Per Johansson：经济学中自然实验和准实验

｜因果科学读书会第5期

【读书会相关信息】

“因果”并不是一个新概念，而是一个已经在多个学科中使用了数十年的分析技术。通过前两季的分
享，我们主要梳理了因果科学在计算机领域的前沿进展。如要融会贯通，我们需要回顾数十年来在社
会学、经济学、医学、生物学等多个领域中，都是使用了什么样的因果模型、以什么样的范式、解决
了什么样的问题。我们还要尝试进行对比和创新，看能否以现在的眼光，用其他的模型，为这些研究
提供新的解决思路。

如果大家对这个读书会感兴趣，欢迎报名：

https://pattern.swarma.org/mobile/study_group/10?from=wechat

【时间】2021年11月28日 19:00--21:00

【主讲人】

Per Johansson，瑞典乌普萨拉大学统计系教授，瑞典皇家科学院院士，诺贝尔经济学奖委员会成
员。1993年毕业于瑞典于默奥大学并获博士学位。Johansson教授致力于微观数据的分析（方法与
应用），研究领域包括劳动经济学（收入保障体系对劳动力市场行为、项目评估、教育经济学的影
响），以及交通经济学等。

个人主页：<https://katalog.uu.se/profile/?id=N99-616>

【邀请嘉宾】

张俊妮，北京大学国家发展研究院副教授。

苗旺，北京大学概率统计系助理教授。

胡悦，清华大学政治学系副教授，博士生导师，清华大学社科学院数据治理研究中心副主任、
Github Campus Advisor。

黄俊铭，普林斯顿大学Associate Research Scientist。

【笔记小分队】

金泽润 北京大学

余枫云 北京协和

【讲座笔记】

研究领域：causal inferences, broadly defined

诺贝尔奖颁给了自然实验以及准实验

Motivation

经济学解决资源分配在人群中的有效分配的问题，兼顾效率与公平。

很多经济学研究，主要是微观计量经济学，关注在于实证性地研究政策对于人和企业的影响。

经济学理论指出一个变量对于另一个变量的因果影响，如教育对于工资的影响。

实证研究的传统，是将理论研究抽象成一个回归方程，进行回归检验。

在实验学科当中，因变量 X 可以被实验者主动地控制，但是现实生活中并非如此，自变量与因变量互相影响，同时还会有第三个变量同时影响自变量和因变量。

线性回归的系数会因为以上的内生性存在而存在偏误。

- Example：人们因为去医院的次数变多而变得健康了吗？调查数据现实去医院次数更多的人反而不健康，然而内生的逻辑是只有不健康的人才会去医院。去医院这件事是一种“选择”，将健康与不健康的人区分开。

应当有一个基于经济学理论的“识别策略”，将自变量的分布外生于实证研究设计，不受研究者控制。

研究一个训练服务对于工资的影响，两个参与主体是政策实施和“你”。加入控制变量并不是“识别策略”。

identification与 estimation之间有区别，多元线性回归是一种estimation而不是identification

RCT将人群区分为实验组和控制组，由于人是智能的，会根据treatment判断其对自己的好处，因此产生了自主选择的偏误。

实验者应当让这种treatment变得具有迷惑性。设计“双盲”，或者“三盲”实验，安慰剂设计是避免这种实验偏误的唯一可能方法。

如果实验的结果是通过问卷调查来得出的，则偏误更严重

RCT的替代性策略是observation study

”准实验“所有观察研究都要有一个设计。Newman Rubin是应用微观经济学的主要框架。

RCM（Rubin Causal Model）是关注treatment的分布，是准实验的一种（一些经济学家有争议）

自然实验，是指对于某个阈值，超过了这个阈值则接受了某种treatment，是某种制度的设计，因此称其为natural experiment

variation也可能被研究者或政策设计者产生。

designed-based 的 observation studies 通常被称为准实验研究。

IV, RD, DID, synthetic control是四种设计方法。

1. 工具变量方法：IV

20世纪30-40年代利用simultaneous equations model来控制变量互相关性。Wright的结构方程的方法，测量供给曲线。

example：劳动的供给需求方程，供给方程测量了一个任意的人对于给定工资，他所工作的时长。但是对于加总的劳动工资和劳动市场数据，对于供给方程来说工资并不是随机分布的。所有所观察到的数据对，都是基于供给和需求方程所得到的均衡，是一系列方程交叉得到的解。应当识别对于供给和需求的一个外生冲击，移动曲线来得到想要估计的曲线。

如果 x_2 包含了 x_1 ，即下方程具有上方程一些没有的变量，则这些变量将被视为”外生“的，可以用作工具变量，通过2SLS进行工具变量回归。

但是这种方法在后来的应用中并不良好，其思想很好但是实证应用比较失败。现在的IV回归更加简洁。

在工具变量回归中的残差，与主回归中的残差是相关的，因此直接进行回归则会得到有偏误的回归参数。

而如果将 x 用工具变量 z 进行拟合替代（第一阶段回归），则会产生无偏的主回归系数，等价于WALD参数（将两阶段回归的回归系数相除）。

example：检验吸烟和婴儿出生体重。随机给婴儿的母亲邮寄吸烟有害的新建。则首先检验在邮寄信件的母亲中多少人停止了吸烟，然后再做主回归。得到戒烟对婴儿出生体重的准确影响，比直接做主回归的影响要大得多。

今日的IV被定义为随机影响treatment的变量。

LATE，工具变量所测度的群体是那些complier，即会受政策变动而摇摆的人，而不是总体。对那些对政策变化不敏感的人没有测度。

总体的影响，应当是得到treatment和没得到的人的加权概率之和。将这个加和概率展开，整理，可以看到将 Y 对 Z 回归可以得到因果效应。整理 α 式子，是将 X 对 Z 回归的结果。

complier指那些受treatment影响的那部分人，如果没收到信件则不会停止抽烟，收到信件就会戒烟。

有批评者认为这种对于complier在人群中的分布并不明确。

如果第一阶段回归的工具变量，对于解释内生变量经济不显著，则会导致非常小的 δ 和非常大的 β ，在这种情况下这个参数并不是正态分布，因此是有问题的。因此需要进行pretesting。

Bayesian analysis是更有效的，可以避免pretesting，是一个较好的补充策略。

2. RD断点回归方法

在几乎半个世纪之前被首创，但是直到1997年才被重新发掘。

Sharp RD

假设一：在连续变量中，定义某一个阈值 c ，如果超过 c ，则视为一种断点。

假设二：产出的条件概率在 c 前后应是连续的。

统计在cutoff周围的产出是否有显著差异。

在实践中，在偏误和效率之间会有权衡。

【精彩问答】

Q1: Casual interpretation of simultaneous model (From Wang Miao)

A: 比如我们已经有了劳动市场数据，从一种equilibrium移动到另一种equilibrium时，比如说鱼的销量与天气变化的关系 这不是动态模型，但是在识别系统模型参数的时候

Q2: Much of the work (Haavelmo, McFadden, Heckman, Card, Angrist, Imbens) in econometrics was paralleled by statistical and biostatistical contributions to the analysis of observational studies, missing data, and case control studies, are statistical contributions to economics or other applied sciences (e.g., medicinal, social, political science) undervalued? (From Wang Miao)

A: 计量更重要的是理论，但是有时理论研究不会看到设计的价值，实证研究有时会改变人们的态度

Q3: 根据我的理解，计量会更接近数学一些，是吗？

A: 确实

Q4: 我更关注这些模型是如何在实际世界中起作用的，机制是非常重要的，请问您是如何理解因果关系机制的

A: 真实世界中异质性的影响是非常大的，Pearl 因果模型能够帮我解释一部分 但是也有很多 variation很难解释

Q5: 您认为Pearl因果模型和Rubin因果模型是从不同角度来解释的还是

A: Pearl的因果图模型，是要求你先假设图的相关关系，再在图的基础上再进行推断，对于我来说Rubin的潜在因果模型在解释误差上更清晰，但是实际上我也认为，这两种模型的运用是否有效取决于具体的情境。从机制视角上，我们通过潜在因果模型能够更加细致的

统计不是数学，是需要理解我们需要如何去应用的学科

（补充by Wang Miao）这两种模型已经被证明在数学上是等价的，你可以从Pearl的图来刻画潜在因果模型，在应用上还是取决于具体的问题背景



pearl1998 Graphs, Causality, and Structural Equation Models.pdf

1.43MB



Graphs, · Causality, and Structural Equation Models.pdf

5.63MB





Q6: 您讨论了很多工具变量 (IV) 方法的细节, 请问您认为它会面临何种问题/挑战呢? IV分析的结果如何指导未来的干预设计?

A: 在一阶段回归中, 渐进分布可能会非常糟糕, 方差可能很大, 当前指导原则需要 () 需要超过100个, Rubin 1997 更好的选择是采用贝叶斯模型 识别哪一组/哪些人比较可能会是compiler, 目前的大家聚焦的主要问题也是在此;

Q7: 您见过贝叶斯方法下的连续工具变量应用吗?

A: 不, 我没有见过, 但是非常有趣, 因为我认为用贝叶斯模型能够比传统统计模型更好地解决这类问题

Q8: 您认为在潜在因果模型领域最重要的开放性问题是什么? 最近在计算科学领域

A: 我之前没有太多考虑过这个问题, 在潜在因果模型中, 我们非常清楚需要去识别的效应, 如果是需要面对连续型干预变量时或者说非常多的干预等级时, 如何来识别效应; 或者如何应用机器学习的方法来识别效应, 我没有太钻研过这个领域, 但是目前我相信目前已经有非常多的相关工作

Q8补充: 现在确实有非常多的工作, 比如决策树, xx来应用于matching等潜在因果模型的

A: (对于这些工作的理解)

Q9 (by Junming Zhang): 如何定量研究工具变量的外生性, 如何解决

A: 敏感性分析, 在阈值附近均匀分布, 通过直方图去看阈值两端协变量是否均衡; 无论是在DID, RD还是在IV设计中, 关键还是在于design时, 让IV提前作用被试 找到一个非常完美的IV是一个非常困难的事情 在我们最近的一篇文章中, 我们在设计时没有结局数据, 这是非常有风险的

(Comment by Junni Zhang) 对于Multiple outcome, 在潜在因果模型中使用matching方法比找到一个合适的工具变量要容易得多

Q10 (by Yanhong Ding): How do you think about the surge of RCTs in managerial practices and entrepreneurship studies those days?

A: 他们能够以非常低的代价重复进行相关的研究

Q11: Would you like to share your ideas about the potential joint point between data-driven machine learning method and econometrics?

A: 机器学习主要还是用于预测, 我认为可以用于Rubin模型中权重的寻找, 但需要我们注意数据的实际情况, 因为实际数据可以非常糟糕

【参考文献】

【读书会PPT】



20211128-Per-Causal Science Reading Clu
b.pdf
1.20MB

