

# 因果推断

潜在结果框架、自然实验与因果图理论 · 博士进阶课程

MarkZZZ WeChat: MarkZZZ20XX

## 课程简介

因果推断 (Causal Inference) 是现代统计学、计量经济学与数据科学的核心方法论, 致力于从观测数据或实验数据中识别**变量间的因果关系**, 而非仅仅描述相关性。本课程系统讲授两大理论框架: 以 Rubin 潜在结果模型 (Potential Outcomes Framework) 为核心的设计类方法, 以及以 Pearl 因果图 (Causal Graphical Models) 为核心的结构方法。课程覆盖从随机对照实验 (RCT)、匹配、工具变量 (IV)、回归断点设计 (RDD)、双重差分 (DID)、合成控制到中介分析、机器学习辅助因果估计 (DML) 的完整工具箱, 并深入讨论每种方法的识别假设、估计理论与推断程序。全部 12 讲以博士水平呈现, 强调**严格的识别理论、完整的数学证明** 以及**前沿方法的最新进展**, 为从事实证研究的学者与数据科学家提供系统的方法论基础。

## 适合人群

- 经济学、统计学、公共政策、社会科学、生物统计等方向的博士生及高年级硕士生
- 希望将因果推断方法应用于产品实验、策略评估的工业界数据科学家
- 对自然实验、政策评估、医疗干预效果评估感兴趣的实证研究人员
- 已掌握基础计量经济学, 希望深入理解识别理论与现代因果方法的研究者

## 前置知识

- **概率统计**: 随机变量、期望、条件期望、大数定律、中心极限定理
- **线性回归**: OLS 估计、Frisch-Waugh-Lovell 定理、异方差稳健标准误
- **线性代数**: 矩阵运算、投影矩阵、特征值分解
- **基础计量经济学 (有益)**: 内生性概念、工具变量直觉、面板数据

## 1 课程内容

| 讲次 | 主题     | 内容概要   |
|----|--------|--|
| 1  | 因果推断框架 | 相关 vs 因果的本质区别; Rubin 潜在结果框架 (Potential Outcomes); SUTVA (稳定单元处理值假设); 平均处理效应 (ATE)、ATT、ATU 的定义与区别; 选择偏差 (Selection Bias) 的分解; 因果推断的根本问题 (Fundamental Problem) |

| 讲次 | 主题          | 内容概要  |
|----|-------------|---|
| 2  | 随机对照实验      | Fisher 精确检验（置换检验完整推导）；Neyman 重复采样框架；有限样本推断与大样本推断；协变量调整（ANCOVA）；分层随机化；双盲设计与溢出效应；实验设计的统计功效分析                            |
| 3  | 可忽略性假设      | 强可忽略性（Strong Ignorability）：无混淆 + 重叠；倾向得分（Propensity Score）的定义与巴拉卡-鲁宾定理完整证明；倾向得分的充分性；重叠（Overlap）假设的必要性；倾向得分估计          |
| 4  | 匹配方法        | 最近邻匹配（1:1, 1:k）；Caliper 匹配；精确匹配与粗化精确匹配（CEM）；逆概率加权（IPW）估计量；增强逆概率加权（AIPW）双重鲁棒估计量的完整理论与证明；匹配估计量的渐近性质                     |
| 5  | 工具变量        | IV 识别框架；二阶段最小二乘（2SLS）的完整推导；弱工具变量问题（F 统计量判别）；LATE（局部平均处理效应）识别定理（Imbens-Angrist）；顺从者、从不顺从者、总是服从者的框架；过度识别检验              |
| 6  | 回归断点设计（RDD） | Sharp RDD 的识别假设与连续性条件；Fuzzy RDD 与 IV 的关系；局部多项式估计与最优带宽（MSE 准则）；Lee-Lemieux 带宽选择；McCrary 密度检验；断点处的平衡性检验；RDD 的实践指南       |
| 7  | 双重差分（DID）   | 平行趋势假设（PT）的形式化与可检验含义；标准 DiD 估计量；事件研究设计（Event Study）；Callaway-Sant'Anna 异质处理效应估计；Sun-Abraham 互动权重估计；Goodman-Bacon 分解定理 |
| 8  | 合成控制法       | Abadie-Diamond-Hainmueller（2010）方法的完整框架；合成控制的权重优化问题；推断方法（置换检验）；合成 DID（SCDiD）；矩阵补全（Matrix Completion）扩展                |
| 9  | Pearl 因果图理论 | 有向无环图（DAG）与因果图； $d$ -分离与条件独立性（Geiger-Verma-Pearl 定理）； $do$ -算子与干预分布；后门准则（完整证明）；前门准则；因果效应的非参数识别                        |

| 讲次 | 主题              | 内容概要  |
|----|-----------------|---|
| 10 | 中介分析            | 直接效应与间接效应的分解；自然直接效应 (NDE) 与自然间接效应 (NIE)；Pearl 中介公式 (完整证明)；Robins-Greenland 反事实框架中的中介；中介分析的识别假设；交叉世界独立性                          |
| 11 | 连续处理与机器学习辅助因果估计 | 广义倾向得分 (GPS)；连续处理的剂量-反应函数 (ADRF)；双重机器学习 (DML/Partially Linear Model) 的 Neyman 正交性原理；R-学习器 (R-Learner)；交叉拟合 (Cross-Fitting) 的必要性 |
| 12 | 前沿专题与研究方向       | 非线性 DID (Callaway-Sant'Anna 2021)；因果发现算法 (PC 算法、FCI 算法)；因果强化学习；大语言模型中的因果推断；研究设计的伦理考量；开放研究问题                                     |

## 2 参考书目

1. Judea Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed., Cambridge University Press, 2009. (因果图理论的奠基之作)
2. Guido W. Imbens and Donald B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences*, Cambridge University Press, 2015. (潜在结果框架的权威教材)
3. Joshua D. Angrist and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009. (自然实验方法的经典入门)
4. Joshua D. Angrist and Jörn-Steffen Pischke, *Mastering 'Metrics: The Path from Cause to Effect*, Princeton University Press, 2014. (计量方法的直观介绍)
5. Scott Cunningham, *Causal Inference: The Mixtape*, Yale University Press, 2021. (现代因果推断方法综述)
6. Nick Huntington-Klein, *The Effect: An Introduction to Research Design and Causality*, CRC Press, 2021. (研究设计视角的因果推断)
7. Alberto Abadie, Alexis Diamond, and Jens Hainmueller, "Synthetic Control Methods for Comparative Case Studies," *Journal of the American Statistical Association*, 105(490), 2010. (合成控制法奠基论文)
8. Guido W. Imbens and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 1994. (LATE 识别定理的原始论文)
9. Victor Chernozhukov et al., "Double/Debiased Machine Learning for Treatment and Structural Parameters," *Econometrics Journal*, 21(1), 2018. (DML 的奠基论文)
10. Brantly Callaway and Pedro H.C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225(2), 2021. (现代 DID 方法的核心论文)
11. James M. Robins and Sander Greenland, "Identifiability and Exchangeability for Direct and Indirect Effects," *Epidemiology*, 3(2), 1992. (中介分析的奠基论文)
12. Judea Pearl, *The Book of Why: The New Science of Cause and Effect*, Basic Books, 2018. (因

果推断的科普与哲学)

MarkZZZ WeChat: MarkZZZ20XX