



# 01反事实因果分析框架



在因果推断中使用机器学习



2023.07.28



纪律：发的材料一律不外传

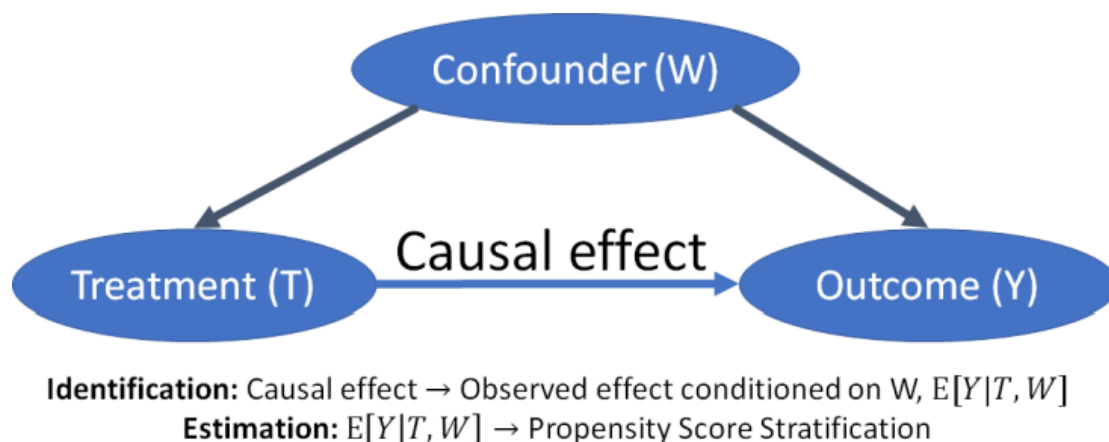
## 1. 反事实因果分析框架

- 项目效应评估（program evaluation）：用于评估政策实施后的效应，**项目效应**也被称为**处理效应**（treatment effect）
- 实验组（treatment group）或处理组（the treated）：项目参与者的全体
- 控制组（control group）或对照组（comparison group）：未参加项目者

### 1.1 以 就业培训 为例

- 考虑就业培训的处理效应，对比实验组和控制组的未来收入及就业情况，发现参加就业培训者的未来收入比未参加者更低

- 分析：是否参加培训是参加者的**自我选择（self selection）**，由于处理组与实验组成员的初始条件不完全相同，故存在**选择偏差**。



## 1.2 反事实框架 (A counterfactual framework)

- 反事实状态：**现实生活中没有发生的状态，响应变量的取值。
- 又称为鲁宾因果模型（Rubin Causal Model, RCM）

If I had left the event early,  
I would not have met  
my soulmate.

Counterfactual

- 记作

$$y_i = \begin{cases} y_{1i} & \text{若 } D_i = 1 \\ y_{0i} & \text{若 } D_i = 0 \end{cases}$$

$$y_i = (1 - D_i) y_{0i} + D_i y_{1i} = y_{0i} + \underbrace{(y_{1i} - y_{0i})}_{\text{处理效应}} D_i$$

- $y_{0i}$  表示  $i$  未参加项目的未来收入， $y_{1i}$  表示  $i$  参加项目的未来收入，个体参加该项目的因果效应为  $(y_{1i} - y_{0i})$
- 处理变量（treatment variable）：**以虚拟变量  $D_i = \{0, 1\}$  表示个体  $i$  是否参与此项目，1 为参与，0 为未参与。
- 结果变量：**感兴趣的结果（outcome of interest）为  $y_i$

## 1.3 ISSUE

- $y_{1i}$  和  $y_{0i}$  不能被同时观测，无法计算个体性的因果关系
- 解决方案：找一些和研究对象 特别像 的个体

## 2. 因果推断的前提假设

### 2.1 个体处理效应稳定假设 (stable unit treatment value assumption, SUTVA)

- 个体处理效应稳定假设：假设样本为 iid, 即对任何  $i \neq j$ ,  $(y_{0i}, y_{1i}, D_i)$  的概率分布与  $(y_{0j}, y_{1j}, D_j)$  相同，而且两者相互独立。这意味着不存在 溢出效应，即个体i是否参与项目，不影响其他任何个体
- SUTVA 假定排除了个体间的 社会互动 (Social interactions) 或者 一般均衡效应 (general equilibrium effects)
- 本质上讲，被研究对象所接受的处理变量的性质是 固定的
  - a. 所关心的处理变量的性质在不同对象之间不存在相互影响
  - b. 处理变量没有版本差异

### 2.2 一致性假设 (consistency)

- 若某个个体接受处理变量影响，则它的观测值就等于它潜在的看不见的那个值

$$(1) \text{ 对于 } D = 1, \quad y_{1i} = y_i | D = 1$$

$$(2) \text{ 对于 } D = 0, \quad y_{0i} = y_i | D = 0$$

### 2.3 可忽略性假设

- 某个个体接受处理是个随机事件，和潜在的看不见的  $y_{0i}$ ,  $y_{1i}$  没有关系

$$D_i \perp (y_{0i}, y_{1i})$$

- 可忽略性假设直接决定了是否有混淆性偏误

### 2.4 正值假设 (positivity)

- 每个被研究对象接受处理变量特定取值的概率都在0到1之间，但不能无限接近0和1

$$0 < P(D_i = 1 | \mathbf{x}_i) < 1$$

- 不能等于0或1是因为：要排除两种分析对象，即  $p=1$  和  $p=0$  的人。因为一些人是不可能接受某些处理变量的，例如假设男生怀孕。

- 当一项研究同时满足正值假设和可忽略性假设时，也称满足“强”可忽略性假设

## 3. 结果的原因和原因的结果

### 3.1 结果的原因

- **问题导向**：先有一个Y，然后倒推回去看什么因素影响了Y，也称**Y导向**
- 在给定需要解释的Y后，Y的潜在解释可能会有很多种，很难确定某个原因有多大的解释力，当确定要研究某个特定变量后，同样不得不面临很多潜在的混淆误差。

### 3.2 原因的结果

- 以原因D为导向，分析重点在于了解D究竟能不能影响Y，影响力有多大
- 目前来说，“**原因的结果**”这个范式在社会科学研究中日渐兴起

## 4. 处理效应与选择难题

### 4.1 处理效应

- 由于处理效应  $(y_{1i} - y_{0i})$  是随机变量，所以我们更关心其期望值
- **参与者平均处理效应 (average treatment effect on the treated, ATT or ATET)**：接受处理变量影响的个体的因果效应。

$$ATT \equiv E(y_{1i} - y_{0i} \mid D_i = 1)$$

- **非参与者平均处理效应 (average treatment effect on the untreated, ATU)**

$$ATU \equiv E(y_{1i} - y_{0i} \mid D_i = 0)$$

- **平均处理效应 (average treatment effect, ATE)**

$$ATE \equiv E(y_{1i} - y_{0i})$$

ATE的计算方法：ATU和ATT的加权平均值（w为相应权重）

$$ATE = ATU * w_{ATU} + ATT * w_{ATT}$$

### 4.2 选择难题

- 在特殊情况下，ATE 与 ATT 可能相等，但一般不相等
- **选择偏差** 问题

$$\begin{aligned}
 & \underbrace{E(y_{1i} | D_i = 1) - E(y_{0i} | D_i = 0)}_{\text{参与者与未参与者的平均差异}} \\
 = & \underbrace{E(y_{1i} | D_i = 1) - E(y_{0i} | D_i = 1)}_{\text{ATT}} + \underbrace{E(y_{0i} | D_i = 1) - E(y_{0i} | D_i = 0)}_{\text{选择偏差}}
 \end{aligned}$$

- **选择难题 (the selection problem)**：由于个体通常会根据其参加项目的预期收益  $E(y_{1i} - y_{0i})$  而自我选择是否参加项目，导致对平均处理效应的估计带来困难。

## 5. 通过随机分组解决选择难题

- 选择难题并非不可克服，解决问题之一就是通过对随机分组 (random assignment) (例如，抛硬币，电脑随机数)，保证了每个研究对象进入到控制组和实验组的概率是一样的。
- 此时

$$E(y_{1i} - y_{0i} | D_i = 1) = E(y_{1i} - y_{0i})$$

- 可得  $ATE = ATT$


$$E(y_{1i} | D_i = 1) - E(y_{0i} | D_i = 0) = E(y_{1i}) - E(y_{0i}) = ATE = ATT$$

- **差额估计量 (difference estimator, 也称difference-in-means, DIM)**：在随机分组的情况下，只需要计算样本中实验组和控制组的平均收入之差，即可一致地估计  $ATE$

$$\hat{DIM} = \frac{\sum_{i=1}^{N_t} Y_i \times D_i}{N_t} - \frac{\sum_{i=1}^{N_c} Y_i \times (1 - D_i)}{N_c}$$

- 均值独立 (mean independence) 只要满足  $y_{0i}$ ， $y_{1i}$  都均值独立于  $D_i$
- 只关心ATT：只需要  $y_{0i}$  均值独立于  $D_i$ ，因为：

$$E(y_{0i} | D_i = 1) - E(y_{0i} | D_i = 0) = E(y_{0i}) - E(y_{0i}) = 0$$

 随机分组可以很好地解决选择偏差，故随机实验或准实验地研究方法在经济学各领域日益盛行

## 6. 依可测变量选择

- 对影响处理对象进行非随机选择地结构变量  $x$  加以控制，可以识别出政策的真正影响。



- 记号：（ $\mathbf{x}$  也称协变量）

$$(y_0, y_1, D, \mathbf{x})$$

- **依可测变量选择 (selection on observable)**：个体  $i$  对  $D_i$  的选择完全取决于可观测的  $\mathbf{x}_i$ ，此时  $(y_{0i}, y_{1i})$  独立于  $D_i$ ，即 **可忽略性假定**。
- **均值可忽略性假定 (Ignorability in Mean)**：在给定  $\mathbf{x}_i$  的情况下， $y_{0i}$  与  $y_{1i}$  都均值独立于  $D_i$ （此假定是通过参数方法和非参数方法一致估计 **ATE**，**ATT** 和 **ATU** 的基础）。

$$E(y_{0i} | \mathbf{x}_i, D_i) = E(y_{0i} | \mathbf{x}_i), \quad E(y_{1i} | \mathbf{x}_i, D_i) = E(y_{1i} | \mathbf{x}_i)$$

- 表示方法

由潜在结果模型

$$\text{当 } D = 1 \text{ 时, } E(Y|\mathbf{x}, D = 1) = E(Y_1|\mathbf{x})$$

$$\text{当 } D = 0 \text{ 时, } E(Y|\mathbf{x}, D = 0) = E(Y_0|\mathbf{x})$$

两式相减

$$E(Y|\mathbf{x}, D = 1) - E(Y|\mathbf{x}, D = 0) = E(Y_1|\mathbf{x}) - E(Y_0|\mathbf{x}) = \text{ATE}(\mathbf{x})$$

又可表达为

$$\text{ATE}(\mathbf{x}) = E(Y|\mathbf{x}, D = 1) - E(Y|\mathbf{x}, D = 0)$$

定义  $m_1(\mathbf{x}) = E(Y|\mathbf{x}, D = 1), \quad m_0(\mathbf{x}) = E(Y|\mathbf{x}, D = 0)$

则有  $\text{ATE}(\mathbf{x}) = m_1(\mathbf{x}) - m_0(\mathbf{x}) = m(\mathbf{x})$

得到

$$\text{ATE} = E_{\mathbf{x}}\{\text{ATE}(\mathbf{x})\} = E_{\mathbf{x}}\{m(\mathbf{x})\}$$

$$\text{ATT} = E_{\mathbf{x}}\{\text{ATE}(\mathbf{x})|D = 1\}$$

$$\text{ATU} = E_{\mathbf{x}}\{\text{ATE}(\mathbf{x})|D = 0\}$$

## 总体 因果参数

$$ATE = E_x\{ATE(\mathbf{x})\} = E_x\{m(\mathbf{x})\}$$

$$ATT = E_x\{ATE(\mathbf{x})|D = 1\}$$

$$ATU = E_x\{ATE(\mathbf{x})|D = 0\}$$

## 样本 因果参数的估计量

$$\widehat{ATE} = \frac{1}{N} \sum_{i=1}^N \hat{m}(\mathbf{x}_i)$$

$$\widehat{ATT} = \frac{1}{\sum_{i=1}^N D_i} \left[ \sum_{i=1}^N D_i \cdot \hat{m}(\mathbf{x}_i) \right]$$

$$\widehat{ATU} = \frac{1}{\sum_{i=1}^N (1 - D_i)} \left[ \sum_{i=1}^N (1 - D_i) \cdot \hat{m}(\mathbf{x}_i) \right]$$

## 7. 依不可测变量选择

- **依不可测变量选择**：项目的选择受与潜在结果相关的不可测变量决定时，条件独立假设或者条件均值独立假设还不够确定平均处理效果。

$$E(Y_1|\mathbf{x}, D) \neq E(Y_1|\mathbf{x}), \quad E(Y_0|\mathbf{x}, D) \neq E(Y_0|\mathbf{x})$$

## 8. 分解选择性偏差

- 将选择性偏差（selection bias）分解为三种可解释的偏差

$$B_I = B_A + B_B + B_C$$

- BA 是由于弱重叠（weak overlap）导致的偏差。
- BB 是由于弱平衡（weak balancing）而存在的偏差。
- BC 是由于存在依不可观测选择（unobservable selection）而存在的偏差。