经济与商务实证研究方法
第一讲

陈志远

人大商学院硕博课程
版本：2025 年春期学期

目录

1 因果推断

1.1 导论

因果关系是科学研究的重要关注对象，是理解客观世界的基础。可以说，实证研究的追求是发现事物之间的因果关系：

- 检验已有理论的真实性帮助预测未来：构建关于现实世界的理论。
- 评估政策的实际效果：税收、补贴、排污权。
- 企业战略、政府政策制定：从复杂的备选项中挑选合理的策略。
- 经济/法律咨询：评价侵权的经济损失。

表 1: 患者康复数据统计

组别	样本	新冠康复时间	标准误
有售卖点的城市	5678	15.3days	1.1
无售卖点的城市	131415	10.1days	2.1

问题 1. 讨论：Bob 公司开发了新冠停特效口服药（kill Covid-19）。如何评估该药的有效性?

- 如何设计实验?
- 要使结果能说明因果关系，应该有哪些关键假定?
- 企业如何根据实验结果进行产品的定价决策? 监管部门如何根据实验结果来进行药物批准上市决策?

假设通过严格实验数据证明了新冠停（kill Covid-19）的有效性，但仅在特定地方售卖，并有如下统计数据:

为什么会发生如此状况?

1.2 因果推断的基础理论

1.2.1 潜在结果分析框架 (Rubin, 1976)

- $D_i = \{0, 1\}$ 表示个体 i 是否接受服药治疗: $D_i = 1$ 表示接受药物治疗，$D_i = 0$ 表示不接受药物治疗;
- 潜在结果（potential outcome）$= \begin{cases} Y_i^1 & \text{if } D_i = 1 \\ Y_i^0 & \text{if } D_i = 0 \end{cases}$，每个个体有两个潜在结果，$Y_i^j, j = 1, 0$ 表示新冠康复时间;
- 实际观测结果为潜在结果的一个:

$$Y_i = Y_i^0(1 - D_i) + Y_i^1 D_i = Y_i^0 + (Y_i^1 - Y_i^0)D_i \tag{1}$$

  ，**其中** $(Y_i^1 - Y_i^0)$ **即为因果效应**;

- 实际观测数据的因果含义：

$$
\begin{aligned}
E[Y_i \mid D_i=1]-E[Y_i \mid D_i=0] &= E[Y_i^1 \mid D_i=1]-E[Y_i^0 \mid D_i=0] \\
&= \underbrace{E[Y_i^1-Y_i^0 \mid D_i=1]}_{\text{ATT}} \\
&\quad+\underbrace{E[Y_i^0 \mid D_i=1]-E[Y_i^0 \mid D_i=0]}_{\text{Selection Bias}}
\end{aligned} \tag{2}
$$

- $E[Y_i^1-Y_i^0 \mid D_i=1]$ 为 ATT(Average treatment effects on the treated），表示在服用新冠停的人群中，新冠停对康复时间的因果效应；
- $E[Y_i^0 \mid D_i=1]-E[Y_i^0 \mid D_i=0]$ 为 selection bias（处理组与对照组的潜在结果不同）

1.2.2 随机试验

在随机试验中，满足：$D_i \perp Y_i^0 \Rightarrow E[Y_i^0 \mid D_i=1]=E[Y_i^0 \mid D_i=0]$ 即没有 Selection bias。随机试验可以解决大部分因果推断的问题，适用于小而美的研究，但缺点在于：(1) 成本高;(2) 结论的普遍意义可能较弱。

注 1. 随机试验是因果分析中的基础思想实验，犹如微观经济学分析中的“完全竞争市场“的感念，帮助提供理想的识别场景。

1.2.3 实验数据的回归分析

- 假设处理效应对所有个体均相同：$Y_i^1-Y_i^0=\rho$
- 可以定义关于平均处理效应的回归方程：

$$
\begin{aligned}
Y_i &= Y_i^0+(Y_i^1-Y_i^0)D_i \\
&= Y_i^0+\rho D_i \\
&= E(Y_i^0)+\rho D_i+Y_i^0-E(Y_i^0)
\end{aligned} \tag{3}
$$

- 定义 $E(Y_i^0)\equiv\alpha$ ，$Y_i^0-E(Y_i^0)\equiv\eta_i$(与均值离差)：

$$
\begin{cases}
E[Y_i^1 \mid D_i=1]=\alpha+\rho+E(\eta_i \mid D_i=1) \\
E[Y_i^1 \mid D_i=0]=\alpha+E(\eta_i \mid D_i=0)
\end{cases}
$$

表 2: 离散随机变量概率分布

$X_i \backslash Y_i$	0	1
0	0.8	0.2
1	0.5	0.5

则有：$E[Y_i^1 \mid D_i = 1] - E[Y_i^1 \mid D_i = 0] = \rho + E(\eta_i \mid D_i = 1) - E(\eta_i \mid D_i = 0)$，如果 η_i 与 D_i 无关，则有：

$$E(\eta_i \mid D_i = 1) - E(\eta_i \mid D_i = 0) = E[Y_i^0 \mid D_i = 1] - E[Y_i^0 \mid D_i = 0]$$

η_i 与 D_i 的相关性即为常见的内生性问题，无选择性偏误保证了无内生性问题，此时回归系数 ρ 具有经济学含义.

问题 2. 当个体处理效应具有异质性时，即 $Y_i^1 - Y_i^0 = \rho_i$ 时，应该如何解释回归方程中 D_i 的系数?

1.2.4 回归分析的统计学理论基础

定义 1. 条件期望方程（Conditional Expectation Function，CEF）

$$E(Y_i \quad | \quad \overrightarrow{x_i} = \overrightarrow{x}) \quad = \int t f_Y(t \mid X_i = x) dt \tag{4}$$

例 1. 给定如下离散随机变量 X_i（是否上大学）与 Y_i（是否高收入）变量的概率分布：

$$\begin{cases} E(Y_i|X_i = 0) = 0.8 \times 0 + 1 \times 0.2 = 0.2 \\ E(Y_i|X_i = 1) = 0.5 \end{cases}$$

注 2. 条件期望是计量分析中非常核心的概念，帮助我们从解释变量学习被解释变量信息。

定义 2. 迭代期望公式（Law of interated expectations）

$$E(Y_i) = E[E(Y_i \mid X_i)] \tag{5}$$

迭代期望公式帮助我们从条件期望变成无条件期望。

表 3: 离散变量概率分布

$X_i \backslash Y_i$	0	1
（0.3）0	0.8	0.2
（0.7）1	0.5	0.5

例 2. 验证迭代期望公式

$$P(x_i = 0) = 0.3P(x_i = 1) = 0.7$$

$$\begin{aligned} E[E(Y_i \mid X_i)] &= 0.3 \times E(Y_i \mid X_i) \\ &= 0 + 0.7 \times E(Y_i \mid X_i = 1) \\ &= 0.3 \times 0.2 + 0.7 \times 0.5 \\ &= 0.06 + 0.35 = 0.41 \end{aligned}$$

$$\begin{aligned} E(Y_i) &= P(Y_i = 0) \times 0 + P(Y_i = 1) \times 1 \\ &= P(Y_i = 1) = 0.3 \times 0.2 + 0.7 \times 0.5 = 0.41 \end{aligned}$$

证明. 连续变量的情形:

$$\begin{aligned} E[E(Y_i \mid X_i)] &= \int E(Y_i \mid X_i = x) \cdot g_x(u) du \\ &= \int [\int t f_Y(t \mid X_i = u) dt] g_x(u) du \\ &= \int t [\int f_Y(t \mid X_i = u) g_x(u) du] dt = E(Y_i) \end{aligned}$$

□

以下是随机变量基于条件期望函数分解的相关定理:

定理 1. *条件期望函数 (Conditional Expectation Function, CEF) 正交分解定理。对于随机变量 Y_i 和 X_i, 有*

$$Y_i = E(Y_i \mid X_i) + \varepsilon_i$$

且满足:(i)$E(\varepsilon_i \mid X_i) = 0$; (ii) $E(\varepsilon_i \cdot g(x_i)) = 0$.

证明. (i)$E(\varepsilon_i \mid X_i)=0$= $E[Y_i-E(Y_i \mid X_i) \mid X_i]$=$E(Y_i \mid X_i)-E(Y_i \mid X_i)=0$; (ii)$E(g(x_i)\cdot\varepsilon_i)$=$E(g(x_i)\cdot E(\varepsilon_i \mid X_i))$=0 □

定理 2. *CEF* 的预测性质

$$E(Y_i \mid X_i)=\underset{m(x_i)}{argmin} E[(Y_i-m(x_i))^2]$$

其中 $m(x_i)$ 是关于 x_i 的任意函数，即$E\left(Y_i|X_i\right)$ 具有最小方差误预测 *MMSE*（*Minimum Mean Squared Error*）

证明. $(Y_i-m(x_i))^2=[(Y_i-E(Y_i \mid X_i))$+$(E(Y_i \mid X_i)$−$m(x_i)]^2$= $(Y_i-E(Y_i \mid X_i))^2+2(E(Y_i \mid X_i)-m(x_i))\times(Y_i-E(Y_i \mid X_i))+(E(Y_i \mid X_i)-m(x_i))^2$ □

定理 3. 方差分解定理。

$$V(Y_i)=V(E(Y_i \mid X_i)+E(V(Y_i \mid X_i))$$

其中 $V(Y_i \mid X_i)=E\left[\left(Y_i-E\left(Y_i|X_i\right)\right)^2|X_i\right]$ 表示**条件方差**（Conditional Variance）。

证明. $Y_i=E(Y_i \mid X_i)+\varepsilon_i$，则有 $V(Y_i)=V(E(Y_i \mid X_i)+V(\varepsilon_i))$=$V(E(Y_i|X_i))+V(\varepsilon_i)+2Cov(E(Y_i|X_i),\varepsilon_i)$，其中

$$Cov\left[E(Y_i|X_i),\varepsilon_i\right]=E\left[\left(E(Y_i|X_i)-E(Y_i)\right)\left(\varepsilon_i-E(\varepsilon_i)\right)\right]$$

其中 $E(\varepsilon_i)=E(E(\varepsilon_i|X_i))=0$, 而根据 CEF 正交分解定理有 $E[E(Y_i|X_i)\varepsilon_i]=0$, 则有 $Cov\left[E(Y_i|X_i),\varepsilon_i\right]=E\left[E\left(Y_i|X_i\right)\varepsilon_i\right]$−$E\left(Y_i\right)E\left(\varepsilon_i\right)=0$. 而 $V(\varepsilon_i)=E[E(\varepsilon_i^2 \mid X_i)]$= $E\left\{E[(Y_i-E(Y_i \mid X_i))^2|X_i]\right\}$=$E(V(Y_i \mid X_i))$ □

注 3. 随机变量的波动程度可以分解为条件期望的波动程度和各条件下方差的期望。