

第 2 章: 無作為化実験による因果効果の推定

Elena Llaudet & Kosuke Imai.

Data Analysis for Social Science: A Friendly and Practical
Introduction.

2026-03-09

2.1 STAR プロジェクト

STAR プロジェクトの概要

- ▶ 1980 年代にテネシー州で実施された「生徒と教師の学習達成率プロジェクト」。
- ▶ 目的: 学級の大きさが生徒の成績に及ぼす影響を検証すること。
- ▶ 幼稚園児を小学 3 年生の終わりまで、以下のいずれかに**無作為に割り当てた**:
 - ▶ 少人数学級 (13 人~17 人)
 - ▶ 標準規模学級 (22 人~25 人)

2.2 処置変数と結果変数

因果関係の方向性

- ▶ **因果関係**: 2つの変数の間の原因と結果の関係。
- ▶ **処置変数 (X)**: 変化が発生する変数 (原因)。
- ▶ **結果変数 (Y)**: 処置変数の変化に応答して変化する変数 (結果)。

$$X \rightarrow Y$$

- ▶ STAR プロジェクトでの関心:
 - ▶ **少人数学級 (処置変数) → 生徒の学業成績 (結果変数)**

処置変数と結果変数の定義

- ▶ **処置変数 (X):**
 - ▶ 処置 ($X_i = 1$): 処置があるという状態 (少人数学級に通った)。
 - ▶ 統制 ($X_i = 0$): 処置がないという状態 (標準規模学級に通った)。
- ▶ **結果変数 (Y):**
 - ▶ reading (読解力テストの点数)
 - ▶ math (算数テストの点数)
 - ▶ graduated (高校卒業を示す二値変数: 1=卒業, 0=卒業しなかった)

2.3 個別因果効果

潜在的結果と個別因果効果

- ▶ **潜在的結果**: 処置を受けた場合と受けなかった場合のそれぞれの結果。
 - ▶ $Y_i(X_i = 1)$: 個人 i が処置を受けた場合の潜在的結果。
 - ▶ $Y_i(X_i = 0)$: 個人 i が統制条件下での潜在的結果。
- ▶ **個別因果効果** = $Y_i(X_i = 1) - Y_i(X_i = 0)$

因果推論の根本問題

- ▶ 現実の世界では、同じ個人について**両方の潜在的結果を同時に観察することは決してない。**
- ▶ 観察できるのは**現実の結果のみ**であり、**反事実の結果**は観察できない。
- ▶ したがって、個人レベルでの因果効果を直接計算することは不可能である。

2.4 平均因果効果

平均因果効果 (平均処置効果)

- ▶ 個人の効果の代わりに、集団における因果効果の平均に注目する。
- ▶ もし全員の両方の潜在的結果を観察できれば：

$$\begin{aligned} & \text{平均処置効果} \\ &= \text{個別因果効果の平均} \\ &= \overline{Y(X = 1)} - \overline{Y(X = 0)} \end{aligned}$$

無作為化実験と平均の差推定量

- ▶ **無作為化実験**: 処置の割り当てが無作為に行われる研究デザイン。
- ▶ 無作為化により、処置群と統制群は、処置以外のすべての特徴において**平均的に互いに同一（比較可能）**になる。
- ▶ この場合、現実の結果を用いて平均処置効果を推定できる：
 - ▶ **平均の差推定量** = $\bar{Y}_{\text{処置群}} - \bar{Y}_{\text{統制群}}$

2.5 少人数学級は生徒の成績を向上させるか?

STAR データの読み込み (1)

▶ 因果効果の推定のため、実験データを R に読み込みます。

1. ローカルに保存した STAR データの読み込み (推奨)

```
star <- read.csv("STAR.csv")
```

(参考) URL から直接読み込むことも可能

```
# star <- read.csv("https://ayumu-tanaka.github.io/QSS/DSS_Data/STAR.csv")
```

データの確認 (2)

```
# データの最初の数行を表示  
head(star)
```

```
##   classtype reading math graduated  
## 1     small     578  610         1  
## 2   regular     612  612         1  
## 3   regular     583  606         1  
## 4     small     661  648         1  
## 5     small     614  636         1  
## 6   regular     610  603         0
```

R の関係演算子と新しい変数の作成

- ▶ == を用いて、特定の値と等しいかを評価できる。
- ▶ ifelse() 関数を用いて、条件に基づいて新しい変数を作成する。

```
# 論理テストの例
```

```
3 == 3
```

```
## [1] TRUE
```

```
3 == 4
```

```
## [1] FALSE
```

```
# 少人数学級に通ったかを示す新しい二値変数 'small' を作成  
star$small <- ifelse(star$classtype == "small", 1, 0)
```

新しい変数 'small' を確認する

▶ データの最初の数行を表示

```
head(star)
```

```
##   classtype  reading  math  graduated  small
## 1     small    578   610           1      1
## 2   regular    612   612           1      0
## 3   regular    583   606           1      0
## 4     small    661   648           1      1
## 5     small    614   636           1      1
## 6   regular    610   603           0      0
```

変数の部分集合化と平均の計算 (1)

▶ [] を用いて、特定の条件を満たす観察を抽出する。

読解力の全体平均

```
mean(star$reading)
```

```
## [1] 628.803
```

処置群 (少人数数学級) の読解力の平均

```
mean(star$reading[star$small == 1])
```

```
## [1] 632.7026
```

変数の部分集合化と平均の計算 (2)

```
# 統制群 (標準規模学級) の読解力の平均  
mean(star$reading[star$small == 0])
```

```
## [1] 625.492
```

平均の差推定量の計算 (読解力と算数)

```
# 読解力についての平均の差推定量
```

```
mean(star$reading[star$small == 1]) - mean(star$reading[star$small == 0])
```

```
## [1] 7.210547
```

```
# 算数についての平均の差推定量
```

```
mean(star$math[star$small == 1]) - mean(star$math[star$small == 0])
```

```
## [1] 5.989905
```

平均の差推定量の計算 (卒業率)

```
# 卒業についての平均の差推定量
```

```
mean(star$graduated[star$small == 1]) - mean(star$graduated[star$small == 0])
```

```
## [1] 0.007031124
```

```
# 解釈の補助
```

```
mean(star$graduated[star$small == 1]) # 処置群の卒業率
```

```
## [1] 0.8735043
```

```
mean(star$graduated[star$small == 0]) # 統制群の卒業率
```

```
## [1] 0.8664731
```

- ▶ **解釈:** 少人数学級に通うことで、高校を卒業する確率が平均で約 0.7 パーcentageポイント上昇した。

2.6 まとめ

第 2 章のまとめ

- ▶ **因果効果**: 処置変数が結果変数に与える影響。
- ▶ **因果推論の根本問題**: 反事実の結果は観察できない。
- ▶ **無作為化実験**: 処置群と統制群を比較可能にし、**平均の差推定量**を用いて平均因果効果を有効に推定できる。
- ▶ R の操作: == 演算子、ifelse() による変数の作成、[] によるデータの部分集合化を学んだ。