統計的因果推論

労働経済学2

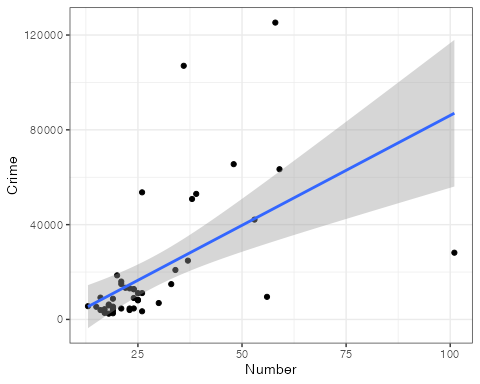
川田恵介

## 統計的因果推論

* 経済理論ととともに、今日的な統計的因果推論の中核的な手法
* ある介入(最低賃金の増額)が、結果（新規就職件数）に与える因果効果
* その他の要因の影響からどう区別するか？
  + 介入の決まり方を推論する (DesignBasedApproach)
  + 結果の決まり方を推論する (OutcomeModelingApproach)

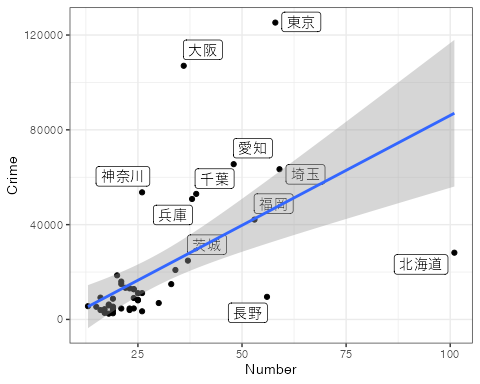
## 例: 図書館 - 犯罪

* 文化施設（図書館）を増やすと、犯罪を減らせるか？



## 例: 図書館 - 犯罪

* 文化施設（図書館）を増やすと、犯罪を減らせるか？



# DesignBasedApproach

* 代表的戦略は、Random Treatment Assignment のもつ性質を用いる
  + ランダム化対照実験: RCT
* 現実世界においてRCTに近い状況を見つける: 自然実験

## Concept: 対照実験

* Controlled Trial
* **同じ**事例を複数用意して、関心のある特徴 (Treatment) のみ変え、関心のある結果 (Outcome) を比べる。
  + 同じビーカーに同じ量の純水を入れ、同じ火力で熱し、沸騰までの時間を比較する。 ただし片方にのみ塩を入れる
  + 差をもたらしうる要因が、Treatmentのみにする
* 多くの応用事例: まったく同じ事例が見つからない | ない | わからない

## ランダム化対照実験(RCT)

* **R**andomized **C**ontrolled **T**rial
* ランダムサンプリングが持つ性質を利用し、近似的な対照実験を行う
* 研究課題・データ収集過程を工夫する

1. 関心となるOutcomeとTreatmentを定める
2. 実験参加者（被験者）を集める
3. Treatmentの水準を**研究者**が被験者間でランダムに決める

## 例: Stats Campの効果

* 大学卒業生・中退者のデータから
  + 留年したかどうか , 入学前の春休みにおける”Stats Camp”
  + Outcome, Treatment
* 関心: Stats Camp参加・不参加者間での留年経験の差

## シンプルなRCT

1. 新入生から”被験者”を集める
2. 被験者を**ランダム**にグループに分割する
3. グループ間の 差を、**母集団における差を推定する手法**で推定する

* グループ間での偶然の偏りは、通常の信頼区間で捉えられる
  + 保守的な結果になる

## 例

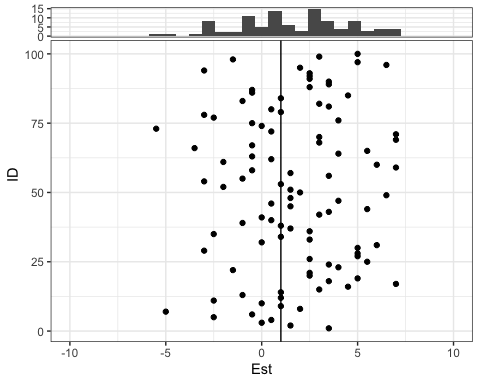
| StataCampに参加した場合 | StataCampに参加しなかった場合 |
| --- | --- |
| 9 | 8 |
| 4 | 3 |
| 7 | 6 |
| 1 | 0 |

* 本当の平均効果

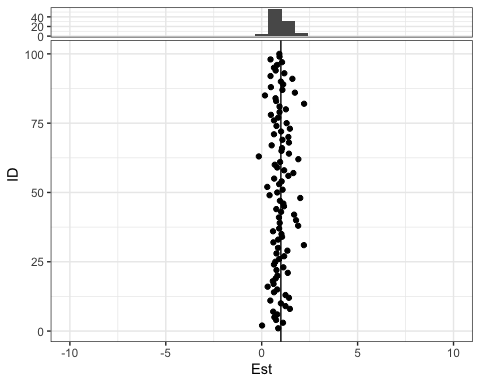
## 例

| StataCampに参加した場合 | StataCampに参加しなかった場合 | D | Y |
| --- | --- | --- | --- |
| 9 | 8 | 1 | 9 |
| 4 | 3 | 1 | 4 |
| 7 | 6 | 0 | 6 |
| 1 | 0 | 0 | 0 |

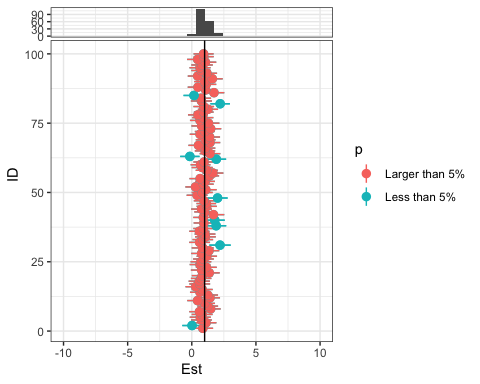
## 例: 参加者 = 4



## 例: 参加者 = 200



## 例: 参加者 = 200



## 自然実験

* 本当に実験できるケースは限られている
* 現実世界においても”条件付き”で実験的な状況は存在する = (自然実験)
  + 同じ内では、介入はランダムに決まっている
* を推定し、 を計算すればOK
  + OLSは一つの選択肢

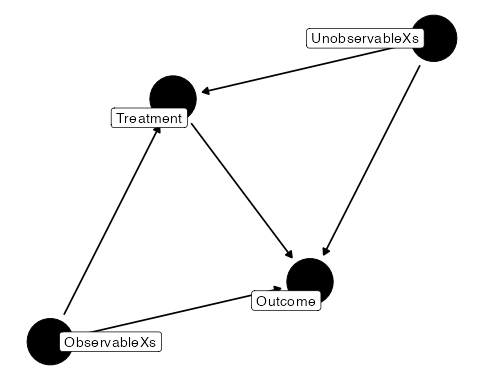
## 例: Life-Event Effect

* 人生経験は、個人の価値観にどのような影響を与えるか？
* 例: 第一子として”女の子”を育てた経験
  + 子供がいる世帯内であれば？

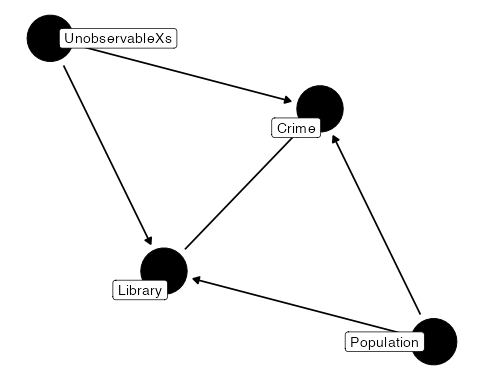
# OutcomeBasedApproach

* のより”正確な”推定を目指す
  + と 双方に影響を与える要因(交絡因子)をモデルに含める

## 概念図



## 観察できる要因



## 観察できる要因

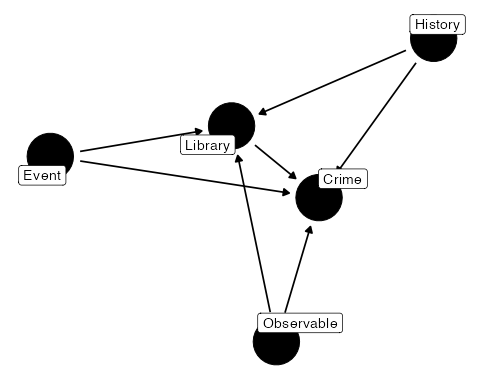
* 推定式に観察できる要因を加える

## 注意点

* 一般にモデルの定式化に推定結果は異存する
* 結果が異なる
* シンプルな対応: 二乗項までは含める
  + 多くの発展的な手法が提案されている

## 観察できない要因

* 一般に観察できない交絡因子が存在する
  + 対応が難しい



## 固定効果

* 観察できない交絡因子「そのものを直接観察できる必要はない」

内で 共通であればいい

* 例:
  + 同一地域であれば、地域の歴史は共通 (地域固定効果)
  + 同一時点であれば、イベントは共通 (時点固定効果)
* ただし まで共通だと比較ができない

## パネルデータ

* 同一地域、個人、企業を継続的に調査
  + 同一主体について、複数時点が観察可能
  + 調査途中でが変化する可能性
* “仮想的に”同一地域、時点内で異なる を比較するモデルを推定できる

## 固定効果モデル

* 以下を推定
* 固定効果モデル

## 注意点

* 地域、時点固有ではない観察できない交絡因子があるとだめ
* モデルの定式化に推定結果が依存する
  + 多くの発展的議論

## まとめ

* 統計的因果推論は学際的に急速な発展を遂げる
* 実務者からも高い関心
  + “経済学部生”に比較優位
* ぜひ継続して勉強してください！！！