

# 研究方法集中セミナー（因果推論）

## 第1回 因果推論の考え方とRCT

---

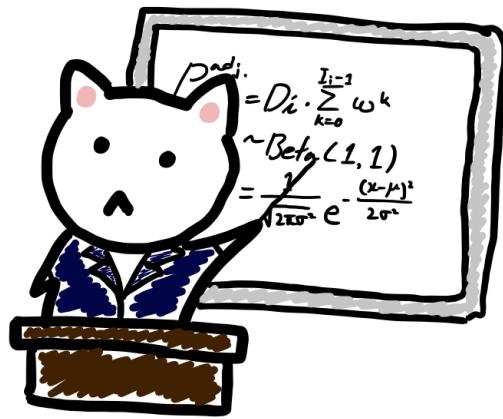
そん じえひよん

久米郁男<sup>1</sup> / 宋 財 泱<sup>2</sup>

<sup>1</sup>早稲田大学政治経済学部 / <sup>2</sup>関西大学総合情報学部

2021年8月5日

# 紹介



LINEスタンプ絶賛販売中!

そん じえひょん

- 宋 財泫 (SONG JAEHYUN)
  - 関西大学総合情報学部
  - 博士 (政治学)
  - 政治行動論、政治学方法論
- [song@kansai-u.ac.jp](mailto:song@kansai-u.ac.jp)
- <https://www.jaysong.net>

# 内容

各講義は以下の内容に関する理論と実習を 5:5 で行う (行いたい)。

## 1. 1日目:8月5日(木)

- 因果推論の考え方
- ランダム化比較試験

## 2. 2日目:8月6日(金)

- Lab Session: R の使い方
- 2日目にLab Sessionを行わない場合は、以下の内容を繰り上げ、5日目は操作変数を解説

## 3. 3日目:8月7日(土)

- 回帰分析とマッチング、その応用

## 4. 4日目:8月9日(月) (オンライン)

- 差分の差分法とその応用

## 5. 5日目:8月10日(火) (オンライン)

- 回帰不連続デザイン

# 実習

実習はRで行う。シラバスで周知した通り、履修者は以下の内容を熟知していると想定するが、1・2日目は(簡単に)Rの使い方についても解説する。

- Jared P. Lander. 2017. *R for Everyone: Advanced Analytics and Graphics* 2nd. Edition. Addison-Wesley Professional
  - 和訳有り(『みんなのR 第2版』)
- 宋財滋・矢内勇生. 『私たちのR: ベストプラクティスの探究』

## 宋のR環境

- OS: macOS Big Sur 11.4
- R 4.1.0 “Camp Pontanezen” + RStudio 1.4.1717
  - R > 4.0ならOK

# 実習

If statistics programs/languages were cars...



# Rの学習資料

## 計量政治学とR

- 浅野正彦・矢内勇生. 2019 『Rによる計量政治学』 オーム社
- 飯田健. 2013. 『計量政治分析』 共立出版

## R全般

- 宋財弦・矢内勇生. 『私たちのR: ベストプラクティスの探究』
  - 無料で読めるハードコアR入門書(執筆中)
- Wickham, Hadley and Grolemund, Garrett. 2017. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, O'Reilly. (邦訳あり/原著はインターネットから無料で閲覧可)
- 松村優哉 他. 2021. 『改訂2版 Rユーザのための RStudio[実践] 入門—tidyverseによるモダンな分析フローの世界—』 技術評論社.
- Wickham, Hadley. 2019. Advanced R, Second Edition, O'Reilly. (邦訳あり/原著はインターネットから無料で閲覧可)

# 評価

## 平常点と期末レポート

- **平常点**: 30%
  - 授業への参加度 (質問/発言)
- **期末レポート**: 70%

# 期末レポート

以下の3つのオプションから選択

## 1. 研究のプロポーザル

- 本講義で紹介した手法を用いた分析のプロポーザルを作成
- 実現可能性があること

## 2. 既存研究の再現 (replication)

1. 分析をそのまま再現
2. 本講義で紹介した手法を用いて再現

## 3. 本講義で紹介した手法を用いた2次分析

# 前提知識

## 本講義の前提知識

- 仮説検定
- 統計的有意性検定
- 回帰分析
- 電卓としてのRの使い方
  - データクリーニング、回帰分析、可視化などができるならベスト
  - 2日目にRの解説は行うが、深入りはしない(できない)
  - 最低限の使い方(データの読み込みなど)は知っている前提で解説
  - 『私たちのR』を読もう

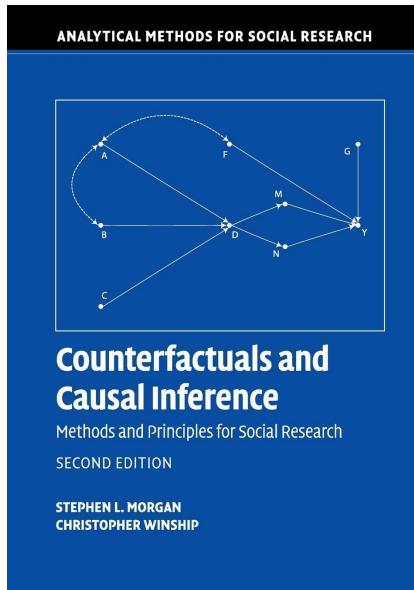
# 因果推論とその敵

---

# 社会科学における因果推論の意味

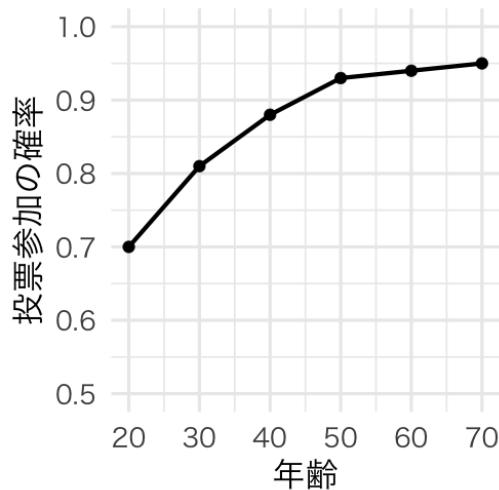
Morgan and Winship (2014) *Counterfactuals and Causal Inference: Methods And Principles For Social Research*. Cambridge

More has been learned about causal inference in the last few decades than sum total of everything that had been learned about it in all prior recorded history. (Gary King)



# 相関から因果へ

## 原因 (X) と結果 (Y) の関係



### 年齢(世代)と投票率の関係 (架空の例)

- 年齢が上がると高い投票率
  - 相関関係
  - 統計分析から得られる結果は相関関係のみ
- 理論/デザインを用いて相関関係が因果関係であることを説得
- 相関関係→因果関係における障害物
  - Simultaneity
  - Spurious Correlation
  - Reverse Causality
  - Omitted Variable Biasなど

# 同時性

## Simultaneity



原因と結果の間に双方向の因果関係が存在

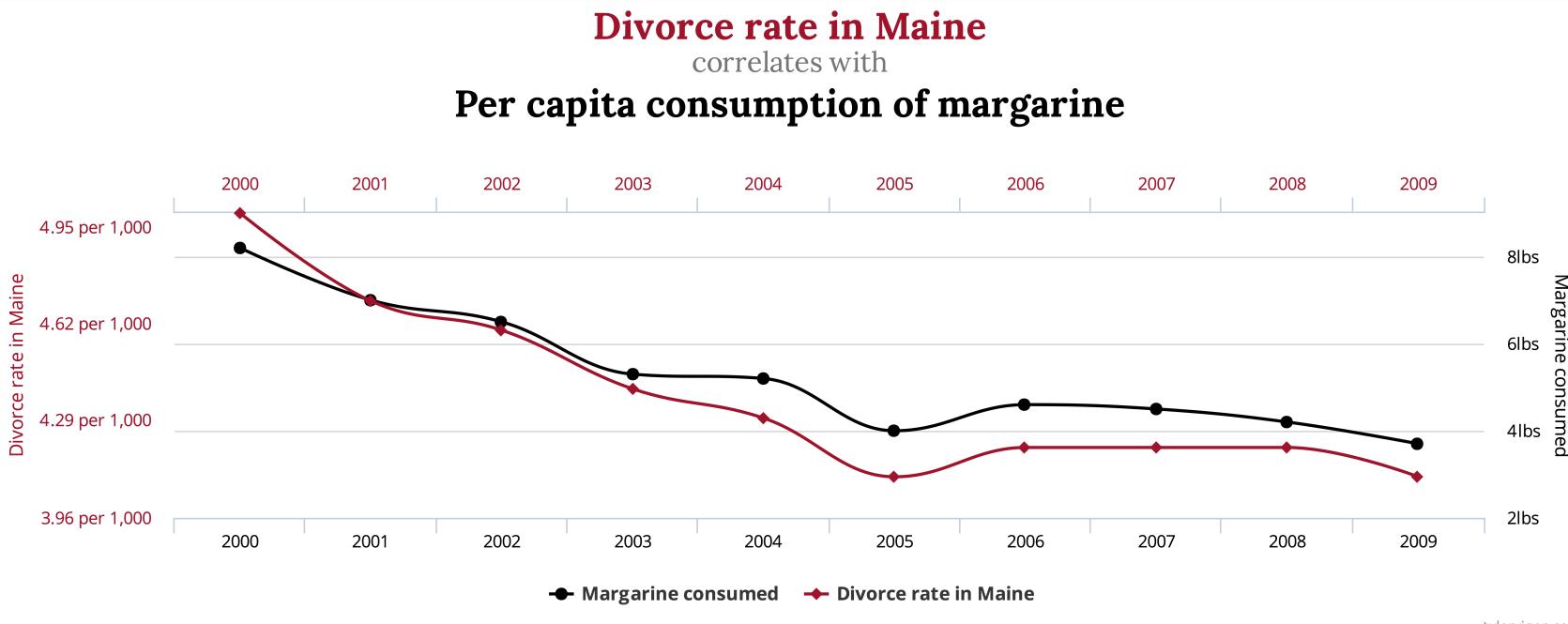
- 例) お酒 (原因; X) とストレス (結果; Y) の関係
  - 酒を飲むとストレスが貯まる
  - ストレス解消のために酒を飲む
  - 酒を飲むとストレスが貯まる
  - ストレス解消のために酒を飲む
  - 酒を飲むとストレスが貯まる
  - ...
- → 地獄のような無限ループ

⇒ 酒がストレスに与える影響は?

# 見かけ上の相関

## Spurious Correlation、擬似相関

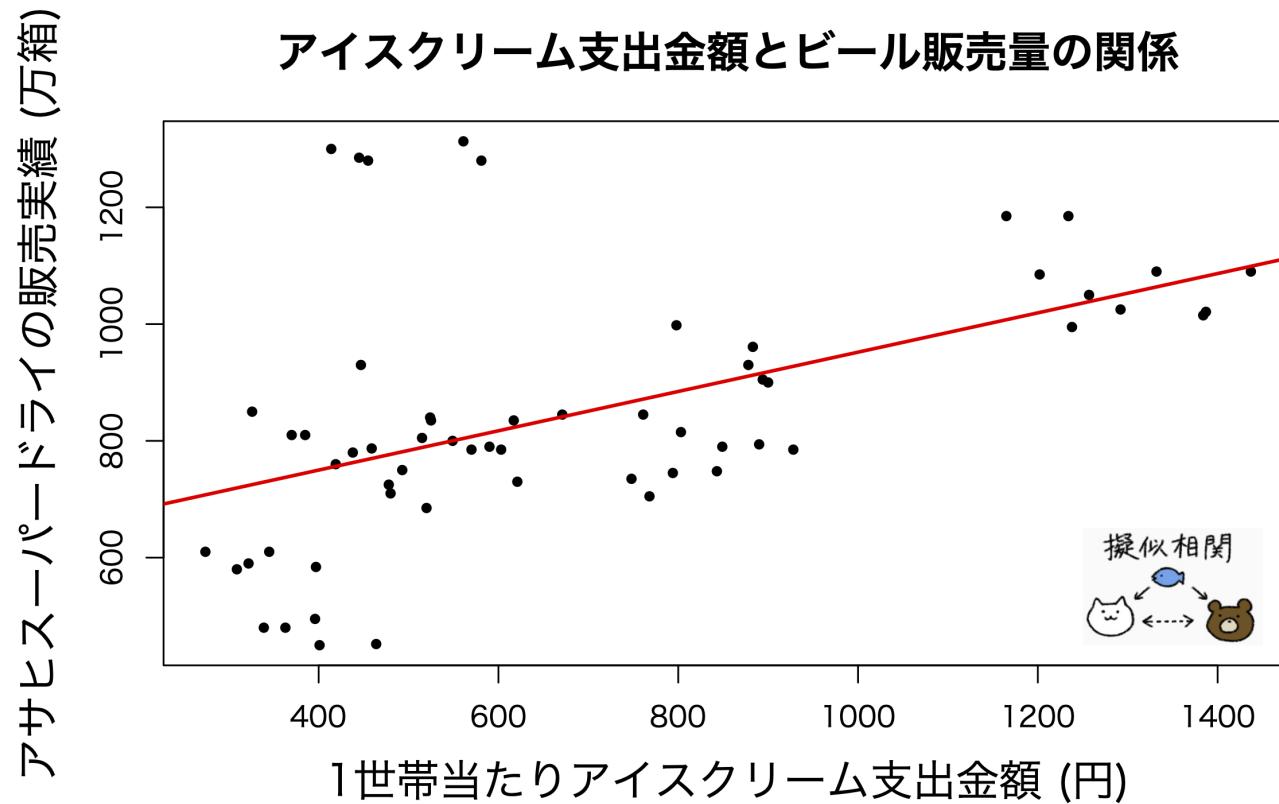
- たまたま相関関係がある場合
  - 例) メイン州の離婚率一人当たりマーガリンの消費量



# 見かけ上の相関

## Spurious Correlation、擬似相関

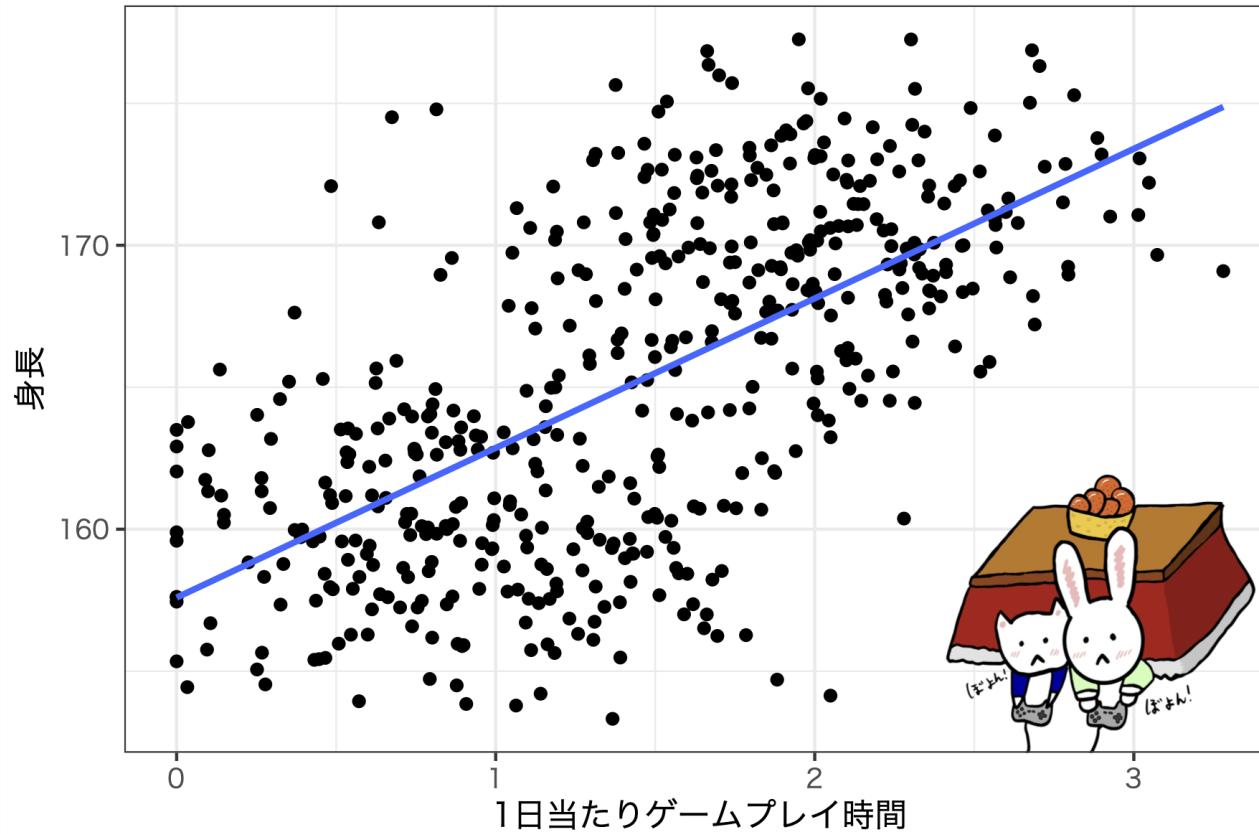
- 共通の要因からの影響
  - 例) ビール消費量とアイスクリーム消費量



# 見かけ上の相関

## Spurious Correlation、擬似相関

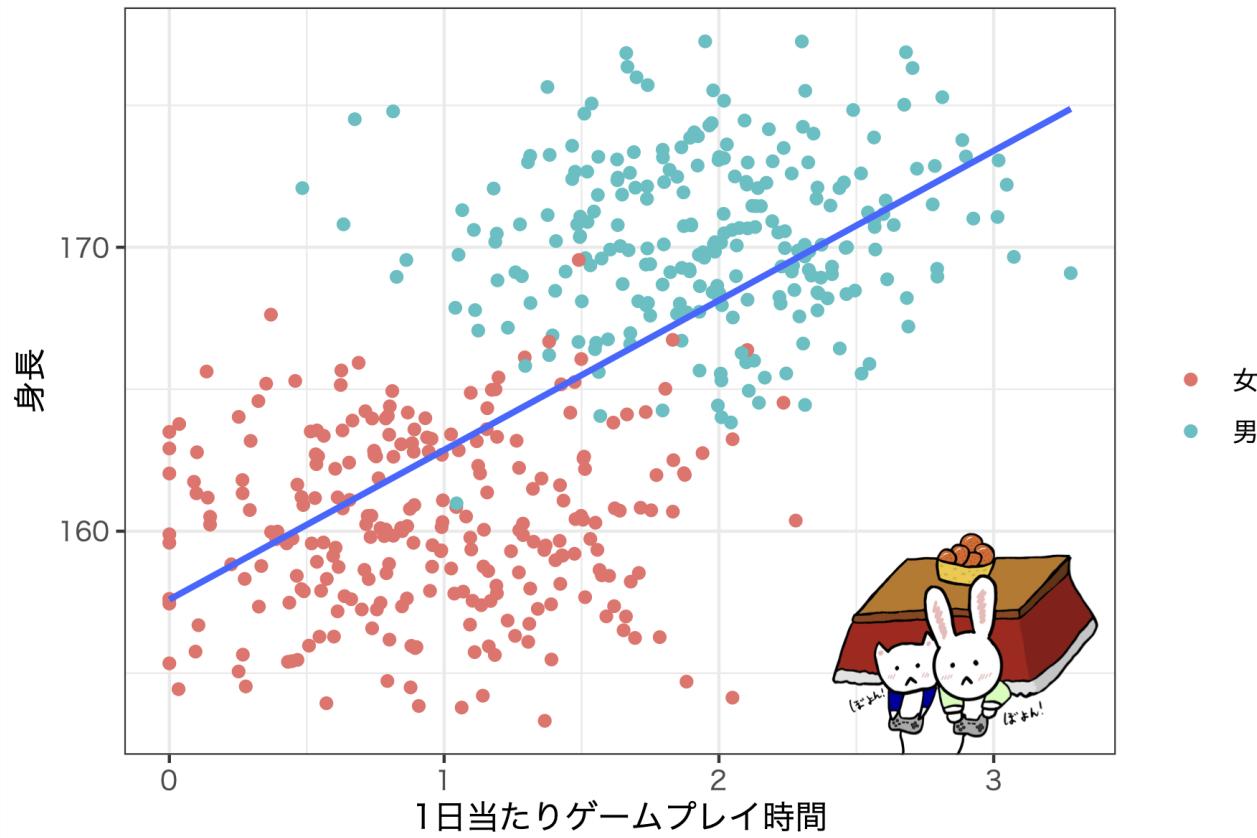
- 共通の要因からの影響
  - 例) ゲームをやると身長が伸びる説



# 見かけ上の相関

## Spurious Correlation、擬似相関

- 共通の要因からの影響
  - 例) ゲームをやると身長が伸びる説



# 逆の因果

## Reverse Causality

- 例) 心臓移植と生存率の例

	5年後に死亡	5年後に生存
心臓移植を受けた	10名	5名
受けなかった	5名	10名

- 心臓移植を受けたら死亡確率が上がる?
- 死亡確率が高い人が心臓移植を受ける?

# 逆の因果

## Reverse Causality

- 「人気だから4文字に略されるのか、4文字に略せるからヒットするのか、どっちなんでしょうね」

多部未華子「わたナギ」も…4文字略称のドラマはなぜ視聴率が伸びる？

8/18(火) 9:26 配信

日刊ゲンダイ DIGITAL



私生活では新婚ほやほやの多部ちゃん (C) 日刊ゲンダイ

「人気だから4文字に略されるのか、4文字に略せるからヒットするのか、どっちなんでしょうね」と話すのは、テレビ誌編集者。

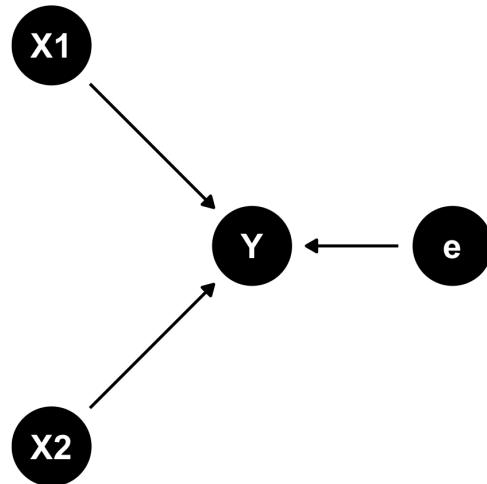
【関連】友人が語る多部未華子の“非モテ時代”

多部未華子（31）主演のドラマ「私の家政夫ナギサさん」（TBS系＝火曜夜10時）もそうだ。

# 欠落変数バイアス

## Omitted Variable Bias

例) 真のモデルが  $Y = \beta_0 + \beta_1 \cdot X1 + \beta_2 \cdot X2 + e$  の場合

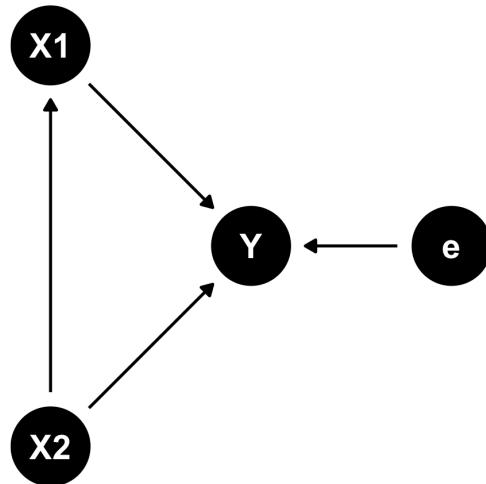


- モデルに  $X2$  が含まれていなくても  $\beta_1$  の推定値は変化×
  - $X_1$  と  $X_2$  が独立
  - $\sigma_{X_1, X_2} = 0$

# 欠落変数バイアス

## Omitted Variable Bias

例) 真のモデルが  $Y = \beta_0 + \beta_1 \cdot X1 + \beta_2 \cdot X2 + e$  の場合



- モデルに  $X2$  が含まれていない場合、 $\beta_1$  の推定値にバイアス
  - $X_1$  と  $X_2$  が**相関**
  - $\sigma_{X_1, X_2} \neq 0$

- 真の  $\beta_1$  を推定するためには  $X1$  と  $Y$  両方と相関する変数**すべて**が必要
  - そもそも、「真のモデル」は未知
- → データ分析から得られた結果はあくまでも「**分析モデルが想定している世界**」のものに過ぎない

# セレクション・バイアス

## Selection Bias

- 例) 職業訓練と期待収入 (参考) 心臓移植

3年後の収入	
職業訓練を受けた	6349ドル
受けなかった	6984ドル

- 職業訓練を受けたら収入が上がる?
- もともと低収入の人が職業訓練を受けようとする?
- 参考) 交絡因子の可能性も (就労意欲など)

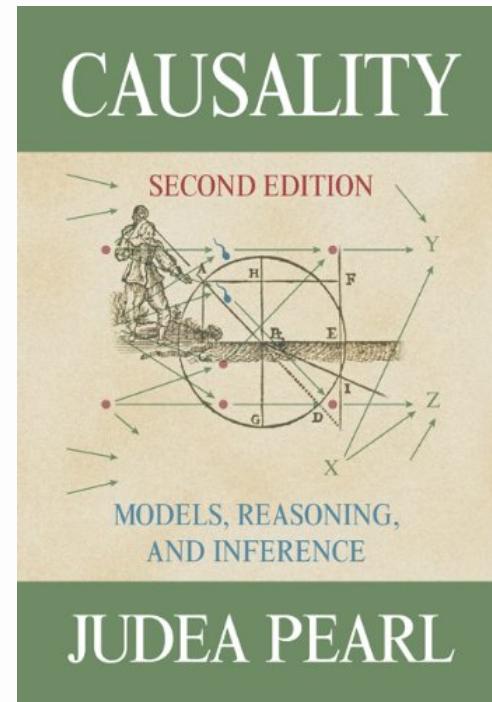
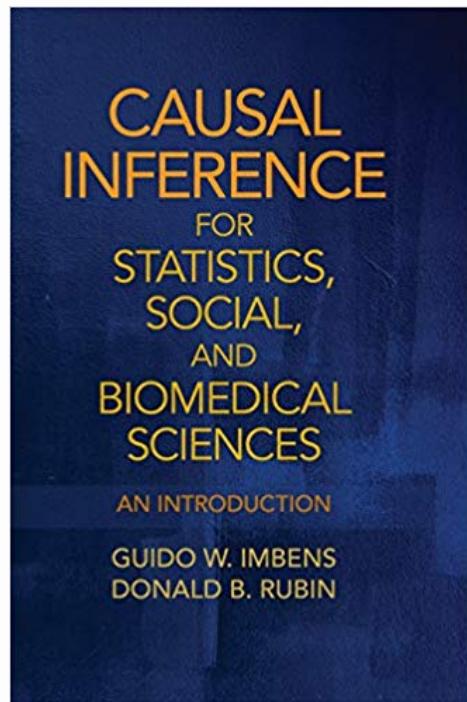
# 内生性

これまでの多くの例は内生性(endogeneity)の問題

- 内生性: 説明変数と誤差項間に相関が存在
  - 誤差項と相関のある説明変数: 内生変数 (endogenous variable)
- 内生性がある場合、推定値は一致推定量ではない
  - $N$  をいくら増やしても無駄無駄無駄
- 内生性の原因
  1. 同時性
  2. 欠落変数バイアス
  3. 測定誤差
  4. セレクション・バイアスなど
- 最近の教科書はこれはすべてをセレクション・バイアスや欠落変数バイアスでまとめる傾向

# 相関から因果へ

内生性の除外 → 因果効果の推定



# 因果推論の根本問題

---

# 因果関係の例

ソンさんの講義を履修することで期待年収が上がるか

- 粕谷君の場合: ソンさんの講義を履修し、年収が5000万円に
  - ソンさんの授業のおかげで富裕層になった(次は社交界進出)
  - 友達に教えてあげよう

講義履修の効果

- 処置: ソンさんの講義を履修するか否か
- 効果: 履修した場合の年収 - 履修しなかった場合の年収

# 因果関係の例

ソンさんの講義を履修することで期待年収が上がるか

- 粕谷君の場合: ソンさんの講義を履修し、年収が5000万円に
  - ソンさんの授業のおかげで富裕層になった(次は社交界進出)
  - 友達に教えてあげよう

講義履修の効果

- 粕谷君がソンさんの授業を履修しなくても年収5000万円なら
  - ソンさんの講義の因果効果は0

履修しなかった場合の年収(A)	履修した場合の年収(B)	効果(B-A)
ケース1	5000万	5000万

# 因果関係の例

ソンさんの講義を履修することで期待年収が上がるか

- 粕谷君の場合: ソンさんの講義を履修し、年収が5000万円に
  - ソンさんの授業のおかげで富裕層になった(次は社交界進出)
  - 友達に教えてあげよう

講義履修の効果

- 粕谷君がソンさんの授業を履修しなくても年収1000万円なら
  - ソンさんの講義の因果効果は4000万円
  - 一生ソンさんには頭が上がらない

---

	履修しなかった場合の年収(A)	履修した場合の年収(B)	効果(B-A)
ケース2	5000万	1000万	4000万

---

# 因果関係の例

ソンさんの講義を履修することで期待年収が上がるか

- 粕谷君の場合: ソンさんの講義を履修し、年収が5000万円に
  - ソンさんの授業のおかげで富裕層になった(次は社交界進出)
  - 友達に教えてあげよう

講義履修の効果

- 粕谷君がソンさんの授業を履修しなくても年収8000万円なら
  - ソンさんの講義の因果効果は-3000万
  - ソンさんは悪くない

	履修しなかった場合の年収(A)	履修した場合の年収(B)	効果(B-A)
ケース3	8000万	5000万	-3000万

# 因果関係の例

ソンさんの講義を履修することで期待年収が上がるか

- 粕谷君の場合: ソンさんの講義を履修し、年収が5000万円に
  - ソンさんの授業のおかげで富裕層になった(次は社交界進出)
  - 友達に教えてあげよう

講義履修の効果

- ソンさんの講義を履修しなかった場合の粕谷君の年収は...?
  - 個人（粕谷君）における処置効果を推定する際にはこれが不可欠

履修しなかった場合の年収(A)	履修した場合の年収(B)	効果(B-A)
ケース1	5000万	5000万 <b>0万</b>
ケース2	5000万	1000万 <b>4000万</b>
ケース3	8000万	5000万 <b>-3000万</b>

# 潜在的結果枠組み

Neyman-Rubin-HollandのPotential Outcome Framework

- $i$ : 学生ID ( $i = 1, 2, 3, \dots, N$ )
- $T$ : 処置
  - 学生  $i$  が謎の薬を飲んだ ( $T_i = 1$ )
  - 学生  $i$  が謎の薬を飲まなかった ( $T_i = 0$ )
- $Y_i(T_i = 1)$ : 学生  $i$  が謎の薬を飲んだ場合の数学成績
- $Y_i(T_i = 0)$ : 学生  $i$  が謎の薬を飲まなかった場合の数学成績
- $ITE_i = Y_i(T_i = 1) - Y_i(T_i = 0)$ : 学生  $i$  における薬の処置効果
  - **ITE**: Individual Treatment Effect (個人における処置効果)
    - = **UTE**: Unit Treatment Effect
  - 全く同じ個人において薬を飲んだ場合と飲まなかった場合の数学成績の差 = 謎の薬の因果効果

# 薬の効果は?

ITEの平均値は-4であり、個人差はあるものの、全体的に薬は成績に負の影響

$i$	$T_i$	$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	$ITE_i$
1	1	77	85	8
2	1	49	59	10
3	1	60	66	6
4	0	61	44	-17
5	0	50	39	-11
6	0	75	55	-20
平均		62	58	-4

# 因果推論の根本問題

しかし、観察できるのは  $Y_i(T_i = 1)$  か  $Y_i(T_i = 0)$ 、片方のみ

- $Y_i(T_i = 0)$  は反実仮想(counterfactual)であり、観察不可 ( $i \in \{1, 2, 3\}$ )
- $Y_i(T_i = 1)$  も反実仮想 ( $i \in \{4, 5, 6\}$ )

$i$	$T_i$	$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	$ITE_i$
1	1	?	85	?
2	1	?	59	?
3	1	?	66	?
4	0	61	?	?
5	0	50	?	?
6	0	75	?	?
平均		62	70	8

「みんなで薬やろうぜ」って言っていいのか

# 世界一受けたいソンさんの授業

$i$	$T_i$	$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	$ITE_i$
1	1	?	700	?
2	1	?	1000	?
3	1	?	550	?
4	1	?	350	?
5	1	?	400	?
6	0	400	?	?
7	0	500	?	?
8	0	350	?	?
9	0	750	?;	?
10	0	500	?	?
平均		500	600	100

# 世界一受けたいソンさんの授業

80万円の価値があるソンさんの講義、みんなで履修しよう!

$i$	$T_i$	$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	$ITE_i$
1	1	550	700	150
2	1	650	1000	350
3	1	600	550	-50
4	1	300	350	50
5	1	300	400	100
6	0	400	300	-100
7	0	500	700	200
8	0	350	600	250
9	0	750	700	-50
10	0	500	400	-100
平均		490	570	80

# 世界一受けたいソンさんの授業

ソンさんは悪くない

$i$	$T_i$	$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	$ITE_i$
1	1	800	700	-100
2	1	650	1000	350
3	1	600	550	-50
4	1	400	350	-50
5	1	350	400	50
6	0	400	300	-100
7	0	500	500	0
8	0	350	400	50
9	0	750	500	-250
10	0	500	400	-100
平均		530	510	-20

# 因果推論の根本問題

- $Y_i(T_i = 1)$  か  $Y_i(T_i = 0)$ 、片方のみしか観察できない状態においてITEから因果効果を推定することは不可能
  - 因果推論の根本問題 (The Fundamental Problem of Causal Inference)
- 解決方法
  - もう一回、過去に戻って異なる処置を行う



# 因果推論の根本問題

- $Y_i(T_i = 1)$  か  $Y_i(T_i = 0)$ 、片方のみしか観察できない状態においてITEから因果効果を推定することは不可能
  - 因果推論の根本問題 (The Fundamental Problem of Causal Inference)
- 潜在的結果を直接観察する方法
  - ただし、個々人の潜在的結果ではなく、集団における潜在的結果
  - 平均処置効果 (ATE; Average Treatment Effect)
  - 無作為割当の重要性

# 平均取るだけでOK?

ただし、観察されたデータから差分を計算するだけではATEは推定不可能

$i$	$T_i$	$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	$ITE_i$
1	1	?	700	?
2	1	?	1000	?
3	1	?	550	?
4	1	?	350	?
5	1	?	400	?
6	0	400	?	?
7	0	500	?	?
8	0	350	?	?
9	0	750	?	?
10	0	500	?	?
平均		500	600	100

# 信頼できるATEの条件

ATE 推定値の信頼性を損なう敵: **内生性** (しかも、常に存在する)

例) やる気のある学生だけがソンさんの講義を履修した場合

## 1. セレクション・バイアス

- ソンさんの講義は鬼畜すぎるため、やる気満々の学生には役に立つものの、やる気のない学生にとってはむしろ学習意欲が低下

## 2. 疑似相関

- やる気のある学生はいろんな方面で頑張るから、将来年収が高くなる。

## 3. 測定誤差

- 履修者の年収はジンバブエ・ドルで測定されている可能性も(これはないか)

内生性は因果推論の敵! どうすれば ...?



**無作為割当** (Random Assignment)

# 無作為割当とは

## 無作為割当 (Random Assignment)

- 処置を受けるかどうかを無作為に割り当てる方法
  - 完全無作為割当: 全ての被験者において、どのグループに属するかの確率が等しい
  - $Pr(T_i = 1) = Pr(T_j = 1)$  where  $i \neq j$
  - $Pr(T_i = 0) = Pr(T_j = 0)$  where  $i \neq j$
  - 無作為割当の方法は色々
- 無作為に割り当てるとき、処置を受けないグループと処置を受けるグループは「集団」として同質なグループになる。
  - 受けないグループ: 統制群 (Control Group)
  - 受けるグループ: 処置群 (Treatment Group)
- 一つの集団を一人の個人として扱い、ITEを測定  $\Rightarrow$  ATE

# 無作為割当の力

コインを投げ、表(  $H$  )なら統制群、裏(  $T$  )なら処置群に割当

- 女性比率が55%、平均年齢が38歳の集団の例

```
SampleData %>%
  summarise(F = mean(Female),
            A = mean(Age))
## # A tibble: 1 x 2
##       F     A
##   <dbl> <dbl>
## 1  0.55   38
```

	ID	Female	Age		ID	Female	Age
	1	1	31		11	0	38
	2	1	41		12	1	29
	3	0	31		13	0	21
	4	1	46		14	0	26
	5	1	37		15	1	36
	6	1	37		16	1	40
	7	0	30		17	0	50
	8	1	46		18	0	42
	9	1	56		19	0	29
	10	0	47		20	1	47

# 無作為割当の力

```
set.seed(19861008)  
Coin <- sample(c("H", "T"), 20, replace = TRUE)  
SampleData$Coin <- Coin
```

ID	Female	Age	Coin	ID	Female	Age	Coin
1	1	31	H	11	0	38	H
2	1	41	T	12	1	29	T
3	0	31	T	13	0	21	H
4	1	46	T	14	0	26	T
5	1	37	H	15	1	36	H
6	1	37	H	16	1	40	T
7	0	30	H	17	0	50	T
8	1	46	T	18	0	42	T
9	1	56	H	19	0	29	H
10	0	47	H	20	1	47	H

# 無作為割当の力

統制群と処置群が比較的同質的なグループに

- 統制群（11名）：女性比率が54.5%、平均年齢が37.2歳
- 処置群（9名）：女性比率が55.6%、平均年齢が39歳

```
SampleData %>%
  group_by(Coin) %>%
  summarise(Female = mean(Female),
            Age     = mean(Age),
            N       = n())
```

```
## # A tibble: 2 x 4
##   Coin   Female     Age     N
##   <chr>   <dbl>   <dbl>   <int>
## 1 H       0.545   37.2     11
## 2 T       0.556   39       9
```

# 無作為割当の力

集団として処置群と統制群は、母集団とほぼ同質

- 母集団:女性率が55%、平均年齢が38歳
- 統制群:女性率が54.5%、平均年齢が37.2歳
- 処置群:女性率が55.6%、平均年齢が39歳
- $n \rightarrow \infty$  なら2つのグループはより同質的に (大数の弱法則)
- 統制群と処置群、母集団はそれぞれ**交換可能** (exchangeable)
  - 処置群に処置を与えること = 母集団全体に処置を与えること
  - 統制群に処置を与えないこと = 母集団全体に処置を与えないこと
- 統制群と処置群の比較で**集団を一つの単位としたITE (=ATE)**が推定可能
  - 処置を与えた母集団 vs. 処置を与えたなかった母集団

# 無作為割当の力

無作為割当は均質な複数のグループを作る手法

- 講義履修と年収の例だと、無作為割当をすることによって ...
  - 各グループにやる気のある学生とない学生が均等に
    - セレクション・バイアス、擬似相関の除去
  - ジンバブエ・ドルで測定される学生も均等に(これはないか)
    - 測定誤差の除去
- 内生性: 处置変数(講義の履修)と誤差項(やる気など)間の相関
  - コイン投げの結果は被験者(学生)の性質と無関係に行われるため、誤差項と相関がない。
  - 外生変数 (Exogenous variable)
  - 学生の性質 ( $X$ ) と处置有無 ( $T$ ) は独立している  $\Rightarrow X \perp T$

無作為割当は内生性を除去する最良の手法

# ランダム化比較試験

---

# ランダム化比較試験とは

## Randomized Controlled Trial (RCT)

- 無作為割当で複数のグループを作り上げた上で、異なる刺激・処置を与え、結果を観察する手法
  - 社会科学でいう「**実験**」の多くはこれを指す
  - **因果推論の王道**
- 因果効果をもたらす(と想定される)処置変数が外生的
  - グループ間における結果変数の差 = 因果効果
- データ生成過程(Data Generating Process; DGP)への直接介入
  - 「**真のモデル**」が分かる

## 実験の方法

- フィールド実験
- 実験室実験
- サーベイ実験など

# データ生成過程への介入

以下のデータ生成過程を仮定

$$\begin{aligned}\text{Income} &= \beta_0 + \beta_1 \cdot \text{Quant} + \varepsilon \\ \varepsilon &\sim \text{Normal}(0, \sigma)\end{aligned}$$

- Income: 10 年後の年収 ( $\in [0, \infty)$ )
- Quant: ソンさんの講義を履修したか否か ( $\in \{0, 1\}$ )
- 誤差項( $\varepsilon$ )には「やる気」や「真面目さ」が含まれるため、Quantと相関がある ( $\rightarrow$  **内生性**)
- 無作為割当て受講有無を決める時、「やる気」や「真面目さ」は Quant と無関係 (**独立**) になる
  - $\Rightarrow$  内生性がなくなる!

# RCTの例

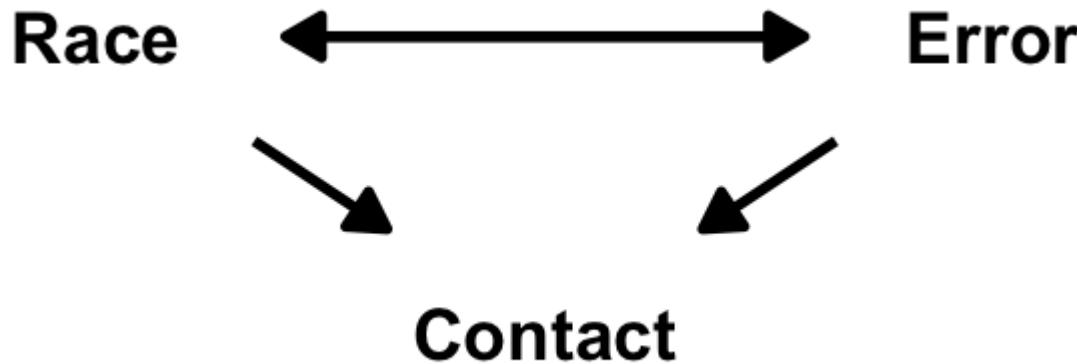
Bertand and Mullainathan (2004)

- 労働市場における人種差別
- 約5000人分の架空の履歴書を求人中の会社へ送る
  - 履歴書の内容 (性別、人種、能力など) は**完全無作為**
  - 履歴書に人種は記入できないため、白人っぽい名前 (Emily など)、黒人っぽい名前 (Jamal など) を記入
- 後は、返事を待つだけ

**処置変数:** 人種 ( $\in \{\text{black, white}\}$ )

**結果変数:** 連絡の有無 ( $\in \{0, 1\}$ )

# 内生性の可能性



- 人種と誤差項には相関があると考えられる
  - $\varepsilon$  には教育水準、親の所得、居住地などが含まれる可能性
  - 実際に人種と上記の要因には相関あり
- 黒人が採用されなかった場合 ...
  - 黒人だから? ← 人種差別 ○
  - 教育水準が低いから ← 人種差別 ×

内生性がある限り、因果効果の識別は不可能

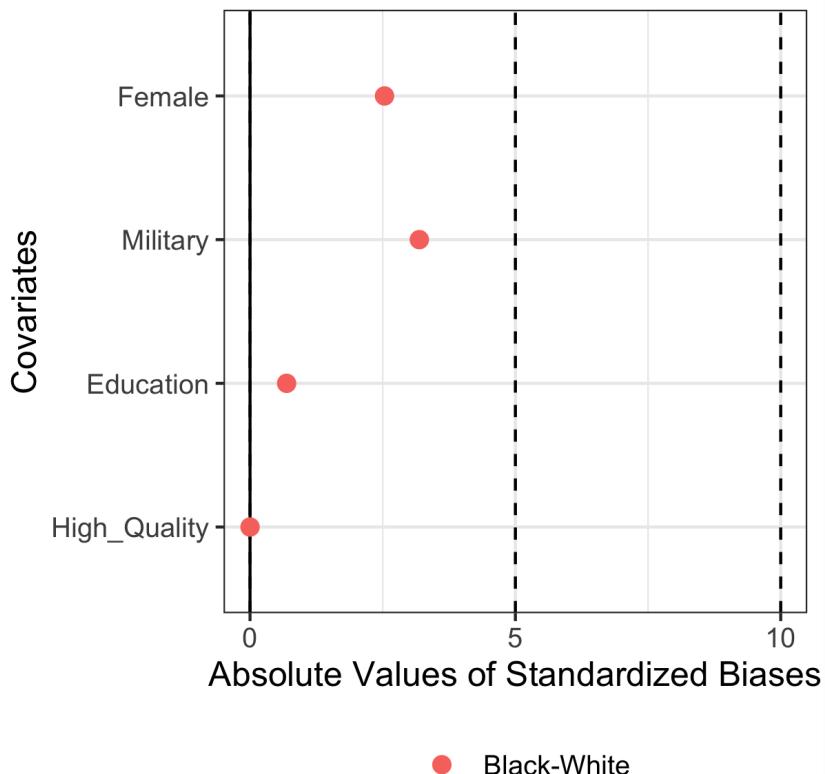
# RCTの力

	白人の名前	黒人の名前
Female	76.42%	77.45%
HighQuality	50.23%	50.23%
Call Rate	9.65%	6.45%
計 (人)	2435	2435

- 無作為割当の結果、人種と性別・能力の相関がほぼ0に
  - 内生性のない状態
  - この場合、労働市場における人種の因果効果は
    - ATE = 黒人の平均連絡率 - 白人の平均連絡率
    - 黒人という理由だけで会社から連絡が来る確率が 3.2%p ↓
    - -3.2%p: 人種の**因果効果** or **処置効果** (treatment effect)

# バランスチェック

無作為割当が行われているか否かを確認



標準化差分を使用

- Standardized Bias (or Difference)
  - サンプルサイズの影響 ×
  - 統計的検定ではない
- $t$  検定、ANOVA、 $\chi^2$  検定は ×
  - バランスチェックに統計的有意性検定は使わない
- `{coblat}`、`{BalanceR}` など

# 標準化差分について

連続変数

$$\text{SB}_{T-C} = 100 \cdot \frac{\bar{X}_T - \bar{X}_C}{\sqrt{0.5 \cdot (s_T^2 + s_C^2)}}$$

二値変数

$$\text{SB}_{T-C} = 100 \cdot \frac{\bar{X}_T - \bar{X}_C}{\sqrt{0.5 \cdot (\bar{X}_T(1 - \bar{X}_T) + \bar{X}_C(1 - \bar{X}_C))}}$$

- $\bar{X}_T$  : 処置群におけるXの平均値
- $s_T^2$  : 処置群におけるXの分散
- $|\text{SB}|$ が小さいほどバランス
- グループが3つ以上の場合、それぞれのペアで実行

# 因果効果の推定

## 方法1: グループ間の結果変数の差分の検定 ( $t$ 検定)

- 因果効果 (ATE):  $\mathbb{E} [\text{Call} | \text{Race} = \text{Black}] - \mathbb{E} [\text{Call} | \text{Race} = \text{White}] = -0.032$
- ATE = 0 の帰無仮説の検定
  - $t_{df=4711.7} = -4.117; p < 0.001; 95\% \text{ CI} = [-0.047, -0.017]$
- 応答変数の尺度に応じてノンパラメトリック分析

## 方法2: 単回帰分析 (線形 or ロジスティック/プロビット)

### 線形回帰分析 (LPM)

Covriates	Est.	S.E.
Intercept	0.064	0.006
Race: White	<b>0.032</b>	0.008

### ロジスティック回帰分析

Covriates	Est.	S.E.
Intercept	-2.675	0.083
Race: White	<b>0.438</b>	0.107

# 因果効果の推定: 重回帰分析は?

無作為割当のおかげですべての変数が互いに独立

- 重回帰分析をしても人種のATEは変化しない (OVBがない)
  - 無作為割当の場合、回帰はしてもしなくても良い
- 現実的に完全にバランスが取れていないため、若干の変化はある

Covariates	Est.	S.E.
Intercept	0.057	0.007
Race: White	0.032	0.08
Female	0.007	0.009
Military	-0.027	0.014
Education	-0.002	0.005
High Quality	0.019	0.008

# 因果効果の不均一性

因果効果が下位グループによって異なる場合

- **因果効果の不均一性** (heterogeneous treatment effects)
  - 例) 性別によって薬の効果が異なる場合
  - 薬の効果が男性なら 1、女性なら 2 の場合
    - 男女比が1:1なら、ATEは1.5に
  - 薬の効果が男性なら 4、女性なら-1 の場合
    - 男女比が1:1なら、ATEは1.5だが...

**方法1:** 男女に分けてATEを推定

**方法2:** 性別と処置有無の交差項を投入した重回帰分析

参考) Bryan, Christopher J., Elizabeth Tipton and David S. Yeager. 2021. "Behavioural science is unlikely to change the world without a heterogeneity revolution," *Nature Human Behaviour*.

# 因果効果の不均一性

## 方法1: 男女に分けてATEを推定

	統制群	処置群	ATE	t	p
男性のみ	0.611	1.561	0.951	-7.521	< 0.001
女性のみ	0.493	2.480	1.987	-15.573	< 0.001
全体	0.551	2.057	1.506	-15.945	< 0.001

```
Data <- read_csv("Data/Day1_Data2.csv")
tibble(Gender = c("Male", "Female"), Female = 0:1) %>%
  mutate(Data = map(Female, ~filter(Data, F == .)),
         ttest = map(Data, ~t.test(Y ~ T, data = .)),
         Fit = map(ttest, broom::tidy)) %>%
  unnest(Fit) %>%
  select(Gender, estimate:p.value, conf.low, conf.high)
```

```
## # A tibble: 2 x 8
##   Gender estimate estimate1 estimate2 statistic  p.value conf.low conf.high
##   <chr>     <dbl>     <dbl>     <dbl>     <dbl>     <dbl>     <dbl>     <dbl>
## 1 Male      -0.951     0.611     1.56     -7.52  1.13e-12    -1.20    -0.702
## 2 Female     -1.99      0.493     2.48     -15.6   4.55e-39    -2.24    -1.74
```

# 因果効果の不均一性

## 方法2: 性別と処置有無の交差項を投入した重回帰分析

```
Fit1 <- lm(Y ~ T * F, data = Data)
stargazer::stargazer(Fit1, type = "html", single.row = TRUE,
                      covariate.labels = c("Treat", "Female", "Treat * Female", "Intercept"),
                      omit.stat=c("n", "f", "rsq", "ser", "adj.rsq"),
                      star.cutoffs = NA, omit.table.layout = "n")
```

<i>Dependent variable:</i>	
Y	
Treat	0.951 (0.131)
Female	-0.117 (0.127)
Treat * Female	1.036 (0.180)
Intercept	0.611 (0.091)

- 男性のATE:  $\beta_1 = 0.951$
- 女性のATE:  $\beta_1 + \beta_3 = 0.951 + 1.036 = 1.987$

# 因果推論の前提:SUTVA

## Stable Unit Treatment Value Assumption

**非干渉性:** 他人の処置・統制有無が処置効果に影響を与えないこと  
例) Aさんの応答変数の値が...

		統制群	処置群
Bさん	統制群	5	10
	処置群	15	20

		統制群	処置群
Bさん	統制群	5	10
	処置群	15	20

**処置の無分散性:** 同じグループに属する対象は同じ処置を受けること

- 手術の場合: 医者、設備、手順、環境など
- 投票参加: 当日、期日前など
- サーベイ実験では問題にならない場合が多い
  - 実験室実験、フィールド実験の場合、「非干渉性」には気をつける
  - 例) 隣の人が見てるのとと私が見てるのが違いますが...?

# 実習

---

# データ

## これまで紹介した例題

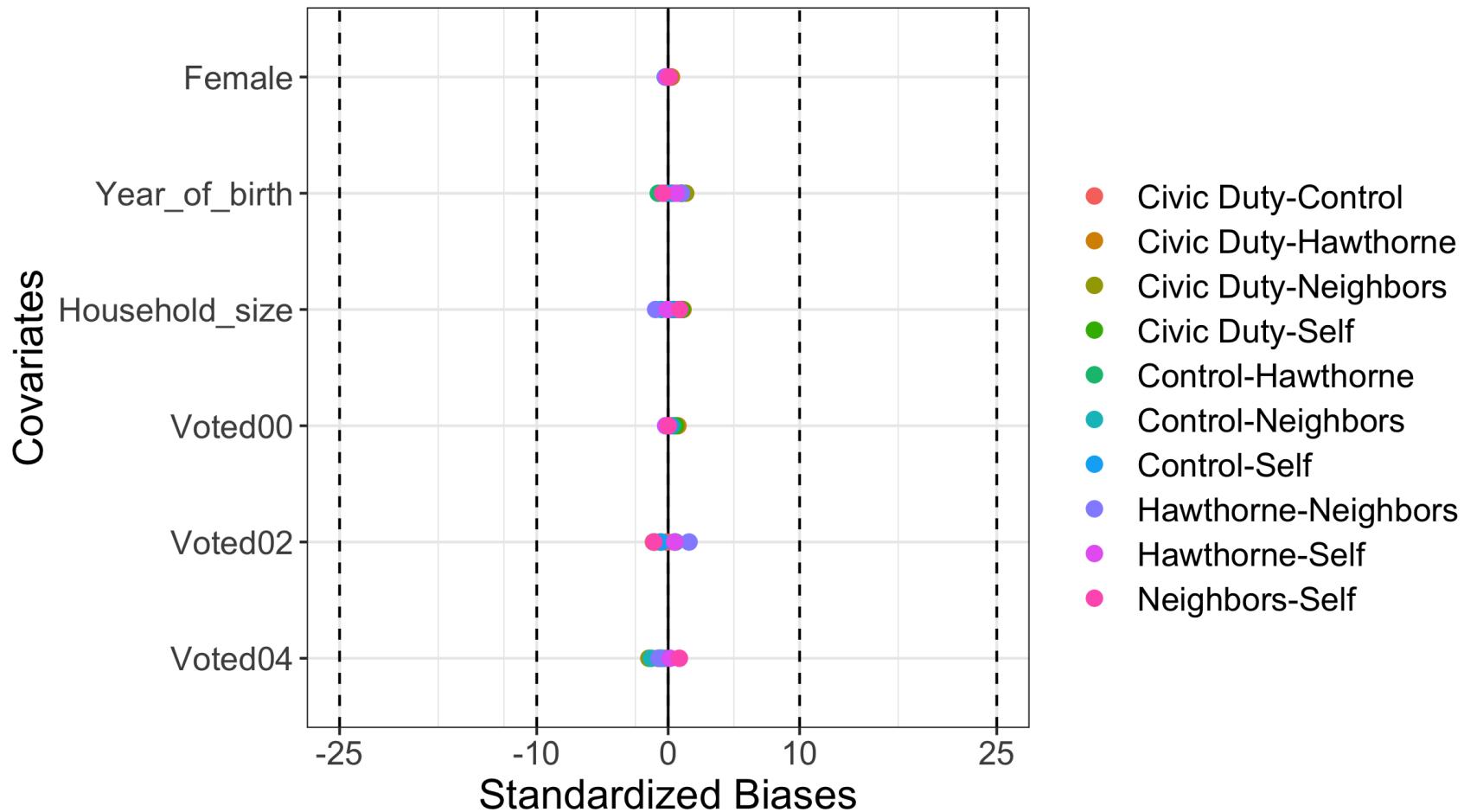
- 労働市場と人種差別: `Day1_Data1.dta`
  - Marianne Bertrand and Sendhil Mullainathan. 2004. “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” *American Economic Review*. 94(4) pp. 991–1013
- 処置効果の不均一性: `Day1_Data2.csv`
  - 架空データ

## Lab Session用のデータ (第2講)

- 社会的圧力と投票参加: `Day1_Data3.csv`
  - Alan S. Gerber, Donald P. Green, and Christopher W. Larimer. 2008. “Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment.” *American Political Science Review*. 102(1) pp. 33–48

# バランスチェック

処置が複数の場合、組み合わせごとに標準化差分を計算



# 処置効果の可視化

