

特殊講義: ヨーロッパ国際政治経済 イントロダクション, RCT

ソンジェヒョン
宋財泫

(SONG Jaehyun)

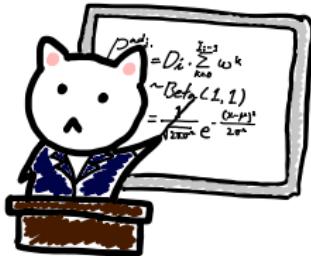
同志社大学文化情報学部 助教

2020年12月23日

講師紹介

そんじえひよん

- 宋財 洋
- 同志社大学文化情報学部 助教
- 政治過程論
 - ▶ 選挙研究、投票行動など
- 政治学方法論
 - ▶ 計量政治学、因果推論、実験政治学など
- jasong@mail.doshisha.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: 12月23日(水)

- ▶ 因果推論の考え方
- ▶ ランダム化比較試験 (Randomized Controlled Trials; RCT)

② 2日目: 12月24日(木)

- ▶ 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: 12月25日(金)

- ▶ 回帰分析とマッチング、その応用
- ▶ クリスマス・パーティーはしない

④ 4日目: 12月26日(土)

- ▶ 差分の差分法とその応用

⑤ 5日目: 12月28日(月)

- ▶ 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』)
- 宋財弦・矢内勇生. 『私たちの R: ベストプラクティスの探究』

講師の R 環境は以下の通り

OS macOS Big Sur 11.0.1

R R 4.0.3 “Bunny-Wunnies Freak Out”

IDE RStudio 1.4.869

- 現在の安定版は 1.3.1093 (2020 年 12 月 10 日)

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

- 宋財滋・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)
<https://www.jaysong.net/RBook/>
- 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. 2018. 『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 50%
 - 授業への参加度 (質問な発言)
 - Slack での参加度も含む
- 期末レポート 50%

Slackについて

- 講義全般に関する連絡は Slack 経由で行う
 - Slack の使い方は Google で「Slack 使い方」で検索
- TA の中島までメール → Slack チャンネルへ招待
 - 既に本講義の Slack スペースに入っている場合は不要
 - タイトルは「因果推論 Slack 登録しろ」 or 「ソンさん、かっこいい！」
 - 空メールで OK

期末レポートについて

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

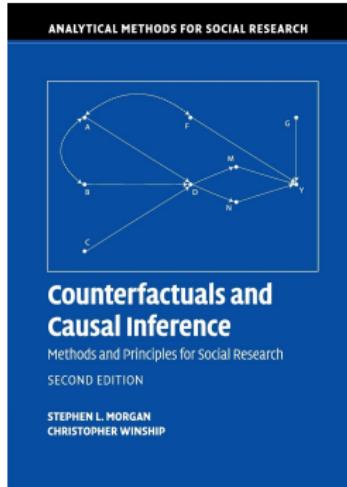
- 浅野正彦・矢内勇生. 2019 『Rによる計量政治学』 オーム社
 - 仮説検定
 - 統計的有意性検定
 - 回帰分析
- 『私たちの R』 第 7~10 章 (できれば 11 章も)
 - 電卓としての R の使い方/データの入出力
 - データ型/データ構造
 - データクリーニング、回帰分析、可視化などができるならベスト
 - 2 日目に R の解説は行うが、深入りはしない (できない)

因果推論とその敵

社会科学における因果推論の意味

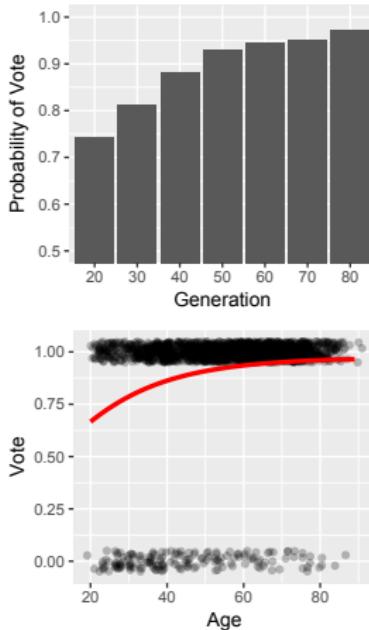
Morgan and Winship (2014) *Counterfactuals and Causal Inference: Methods And Principles For Social Research*. Cambridge

“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

- 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
 - 他にも色々

因果推論の敵 (1) Simultaneity

同時性

- 原因と結果の間に双方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
 - ...

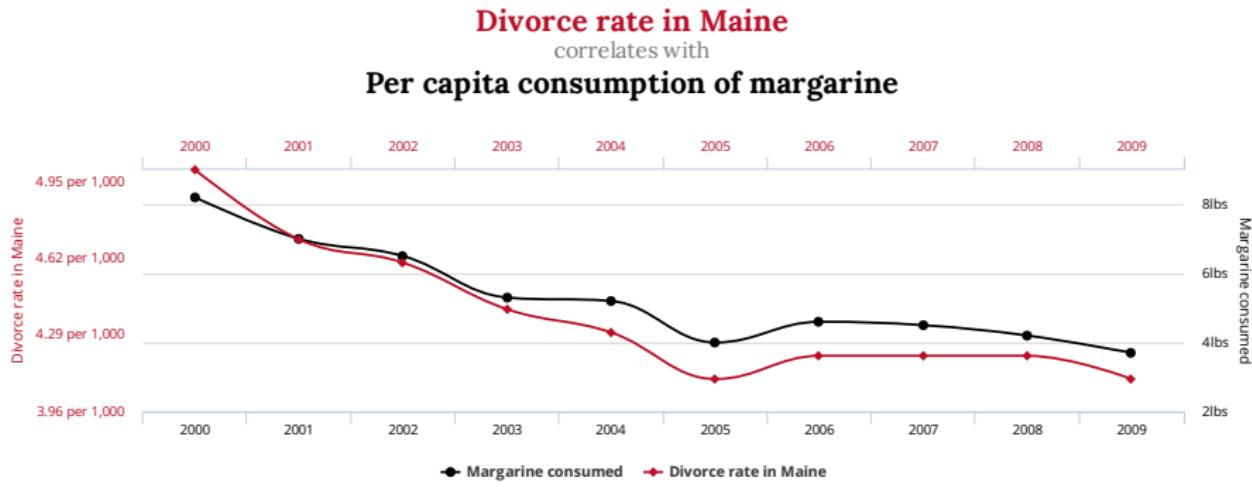
→ 地獄のような無限ループ

⇒ 酒がストレスに与える影響は?



因果推論の敵 (2) Spurious Correlation

疑似相関

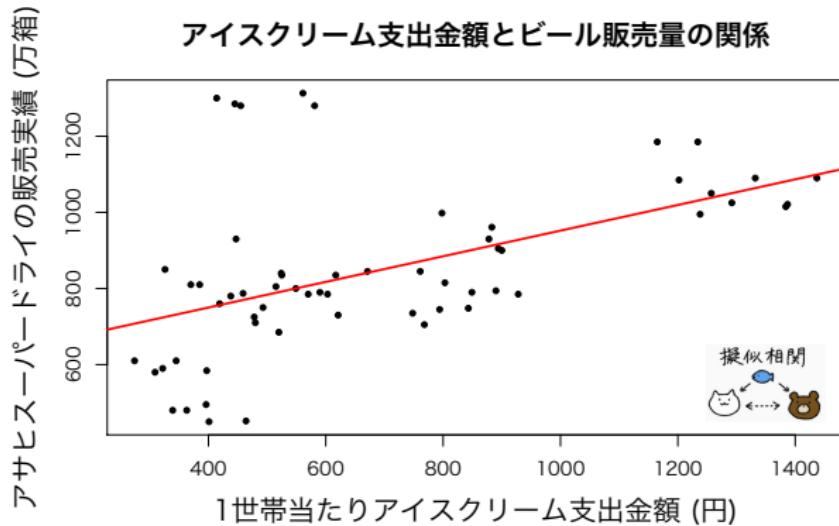


① たまたま相関関係がある場合

例) メイン州の離婚率一人当たりマーガリンの消費量

因果推論の敵 (2) Spurious Correlation

疑似相関

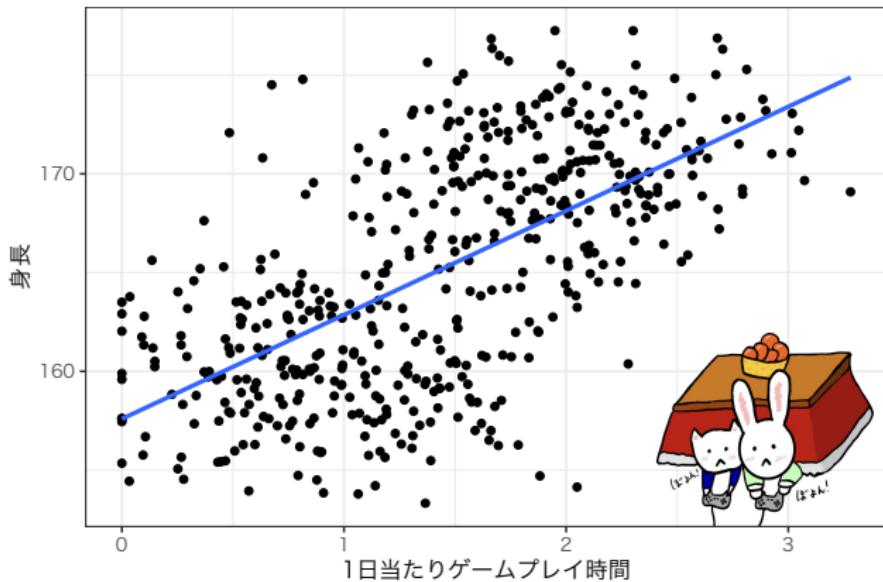


② 共通要因からの影響

例) ビール消費量とアイスクリーム消費量

因果推論の敵 (2) Spurious Correlation

疑似相関

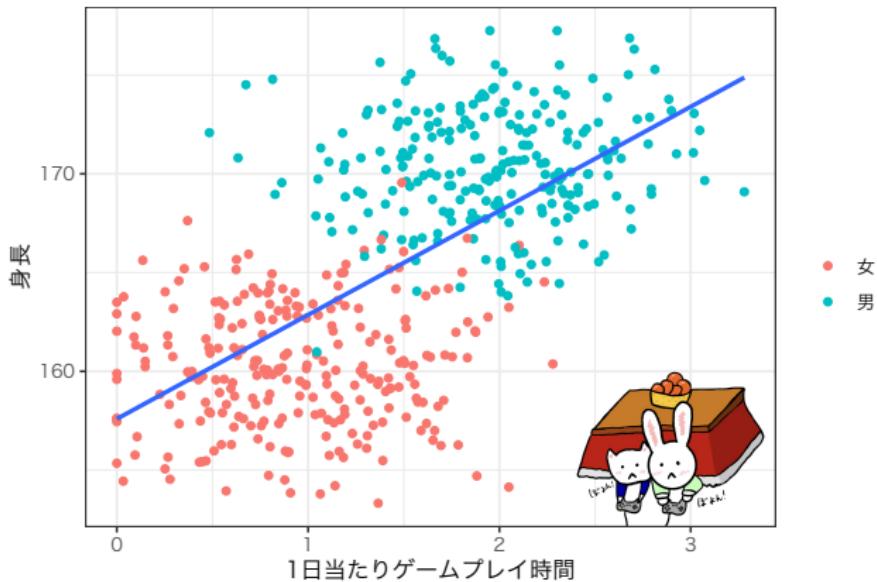


② 共通要因からの影響

例) ゲームをやると身長が伸びる説

因果推論の敵 (2) Spurious Correlation

疑似相関



② 共通要因からの影響

例) ゲームをやると身長が伸びる説

因果推論の敵 (3) Reverse Causality

逆の因果関係

心臓移植と生存率の例

		5年後に	
		死亡	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率が上がる?
- ⇒ 死亡リスクの高い人に心臓移植を施術する?

因果推論の敵 (3) Reverse Causality

逆の因果関係

- 「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。

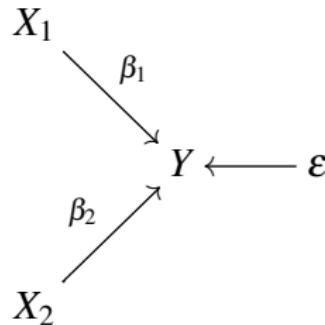
【関連】友人が語る多部未華子の“非モテ時代”

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。

因果推論の敵 (4) Omitted Variable Bias

欠落変数バイアス

真のモデルが $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ の場合



- モデルに X_2 が含まれていなくても β_1 の推定量は変化 ×
= X_1 と X_2 が独立
= $\text{Cov}(X_1, X_2) = 0$

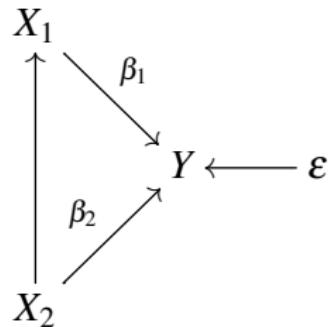
真の β_1 を推定するためには X_1 と Y 両方と相関する変数すべてが必要

- 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

因果推論の敵 (4) Omitted Variable Bias

欠落変数バイアス

真のモデルが $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ の場合



- X_2 を投入しないと、 β_1 の推定量が変化
 - $X_1 \rightarrow Y$ の影響を推定する際に X_2 の影響は無視できない
- = X_1 と X_2 が連関
- = $\text{Cov}(X_1, X_2) \neq 0$
- ⇒ 欠落変数バイアス

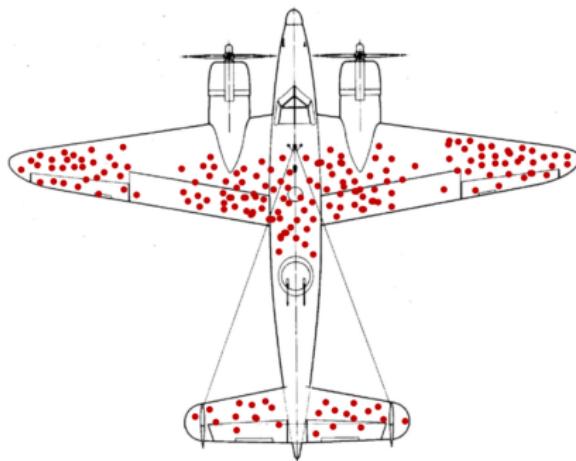
真の β_1 を推定するためには X_1 と Y 両方と相関する変数すべてが必要

- 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

因果推論の敵 (5) Selection Bias

選択バイアス

- 主翼に多くの被弾の痕跡
- 主翼を鉄で補強すれば、より頑丈になる?
 - エンジンは丈夫だから被弾の痕跡がない?



因果推論の敵: Endogeneity

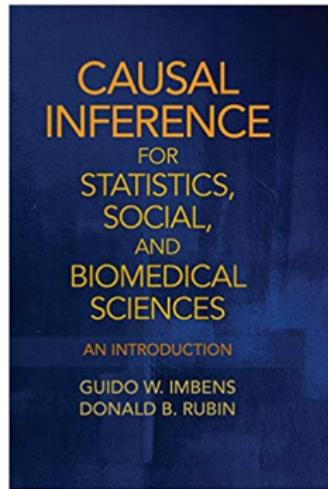
これまでの多くの例は内生性 (Endogeneity) の問題

- 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
 - 信頼できる推定量ではない
- 内生性の原因
 - ① Simultaneity
 - ② Omitted Variable Bias
 - ③ Measurement Error
 - ④ Selection Bias などなど
 - これらは数学的に非常に似ている場合が多く、あるバイアスを他のバイアスとして解釈することも可能
 - 最近の教科書はこれはすべてをセレクションバイアスでまとめたり、欠落変数バイアスでまとめる傾向

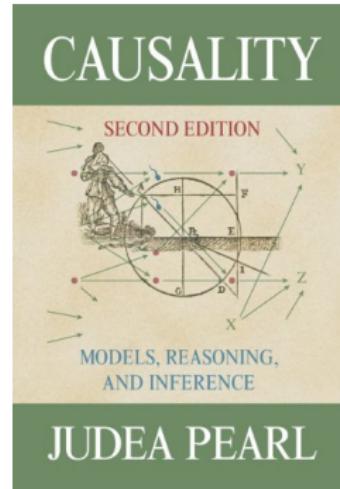
相関から因果へ

因果推論の2大勢力

- 「内生性の除外 → 因果効果の推定」という目的は同じ



潜在的結果アプローチ



構造的因果モデル

因果推論と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

- 中島の場合：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

講義の履修の効果

- 処置：ソンさんの講義を履修するか否か
- ITE : 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

- 中島の場合：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

講義の履修の効果

- 中島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

	年収 (万円)	ITE (万円)
	履修 × (A)	履修 ○ (B)
ケース 1	5000	5000
		0

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

- 中島の場合：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

講義の履修の効果

- 中島がソンさんの授業を履修しかった場合の年収が 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

	年収 (万円)		ITE (万円) (B - A)
	履修 × (A)	履修 ○ (B)	
ケース 2	1000	5000	4000

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

- 中島の場合：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

講義の履修の効果

- 中島がソンさんの授業を履修しかった場合の年収が 8000 万円なら
 - ⇒ ソンさんの講義の因果効果は -3000 万円
 - ⇒ ソンさんは悪くない

	年収 (万円)		ITE (万円) (B - A)
	履修 × (A)	履修 ○ (B)	
ケース 3	8000	5000	-3000

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

- 中島の場合：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

講義の履修の効果

- ソンさんの講義を履修しなかった場合の中島の年収は ... ?
 - ITE の推定の際にはこれが不可欠

	年収 (万円)		ITE (万円) (B - A)
	履修 × (A)	履修 ○ (B)	
ケース 1	5000	5000	0
ケース 2	1000	5000	4000
ケース 3	8000	5000	-3000

Potential Outcome Framework

Neyman-Rubin-Holland の「潜在的結果枠組み」

- 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
- $\text{ITE}_i = Y_i(T_i = 1) - Y_i(T_i = 0)$: 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)
 - 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**

Potential Outcome Framework

Neyman-Rubin-Holland の「潜在的結果枠組み」

i	T_i	Y_i		ITE $_i$
		$\bar{Y}_i(T_i = 0)$	$\bar{Y}_i(T_i = 1)$	
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

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

The Fundamental Problem of Causal Inference

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
 - $Y_{i \in \{1,2,3\}}(T_{i \in \{1,2,3\}} = 0)$ は**反実仮想** (counterfactual) であり、観察不可
 - $Y_{i \in \{4,5,6\}}(T_{i \in \{4,5,6\}} = 1)$ も**反実仮想**

i	T_i	Y_i		ITE _i
		$Y_i(T_i = 0)$	$Y_i(T_i = 1)$	
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

The Fundamental Problem of Causal Inference

世界一受けたいソンさん授業の例

学生 ID	履修	10 年後の年収 (万円)	
		履修しない場合	履修する場合
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

The Fundamental Problem of Causal Inference

世界一受けたいソンさん授業の例

学生 ID	履修	10 年後の年収 (万円)	
		履修しない場合	履修する場合
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

The Fundamental Problem of Causal Inference

世界一受けたいソンさん授業の例

学生 ID	履修	10 年後の年収 (万円)			ITE
		履修 ×	履修 ○		
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	

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！

The Fundamental Problem of Causal Inference

世界一受けたくないソンさん授業の例

学生 ID	履修	10 年後の年収 (万円)			ITE
		履修 ×	履修 ○		
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	

今すぐ履修をやめて、有意義な時間を送ろう！

The Fundamental Problem of Causal Inference

- $Y_i(T_i = 1)$ 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)

The Fundamental Problem of Causal Inference

- $Y_i(T_i = 1)$ 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



The Fundamental Problem of Causal Inference

- $Y_i(T_i = 1)$ 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できる方法がある
 - ただし、個々人の因果効果ではなく、集団における平均的な因果効果
 - 平均処置効果 (ATE; Average Treatment Effect)
 - 無作為割当の重要性

The Fundamental Problem of Causal Inference

単に、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

学生 ID	履修	10 年後の年収 (万円)	
		履修しない場合	履修する場合
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

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
 - 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 $i \neq j$
 - 条件付き無作為割当というものもある
- 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

- 女性率 60%、平均年齢 38.25 歳の集団の例

ID	性別	年齢	コイン	ID	性別	年齢	コイン
1	女	20		11	男	45	
2	女	31		12	男	55	
3	男	27		13	女	25	
4	男	21		14	男	43	
5	女	44		15	女	47	
6	女	44		16	女	26	
7	女	39		17	女	23	
8	女	46		18	男	29	
9	男	59		19	女	41	
10	女	58		20	男	42	

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

統制群 女性率 60%、平均年齢 37.7 歳

処置群 女性率 60%、平均年齢 38.8 歳

ID	性別	年齢	コイン	ID	性別	年齢	コイン
1	女	20	<i>T</i>	11	男	45	<i>T</i>
2	女	31	<i>H</i>	12	男	55	<i>H</i>
3	男	27	<i>H</i>	13	女	25	<i>T</i>
4	男	21	<i>H</i>	14	男	43	<i>T</i>
5	女	44	<i>T</i>	15	女	47	<i>H</i>
6	女	44	<i>T</i>	16	女	26	<i>H</i>
7	女	39	<i>T</i>	17	女	23	<i>H</i>
8	女	46	<i>T</i>	18	男	29	<i>T</i>
9	男	59	<i>H</i>	19	女	41	<i>H</i>
10	女	58	<i>H</i>	20	男	42	<i>T</i>

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 60%、平均年齢が 38.25 歳
 - 統制群：女性率が 60%、平均年齢が 37.70 歳
 - 処置群：女性率が 60%、平均年齢が 37.80 歳
 - $N = \infty$ の場合、統制群と処置群は母集団と完全に一致 (大数の法則)
-
- 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 ATE が推定可能
- = 処置を与えた母集団 vs. 処置を与えなかった母集団

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：処置変数 (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。
 - このような変数を **外生変数** (Exogenous variable) と呼ぶ
 - 学生の性質 (X) と処置有無 (T) は独立している $\Rightarrow X \perp T$ と表記する

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

RCT とは？

ランダム化比較試験 (Randomized Controlled Trial; RCT)

- 無作為割当て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 (DGP: Data Generating Process) への直接介入
 - 「真のモデル」 が分かる ← 超重要

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

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

- 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 (ε) 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と、「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

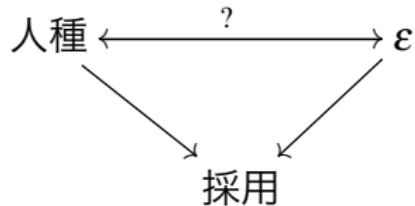
Bertand and Mullainathan (2004)

- 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) は完全ランダム
 - 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 ($\in \{\text{black}, \text{white}\}$)

結果変数 連絡の有無 ($\in \{0, 1\}$)

内生性の恐れ



- 人種と誤差項には相関があると考えられる
 - ϵ 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 ...
 - ① 黒人だから? ← 人種差別 ○
 - ② 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

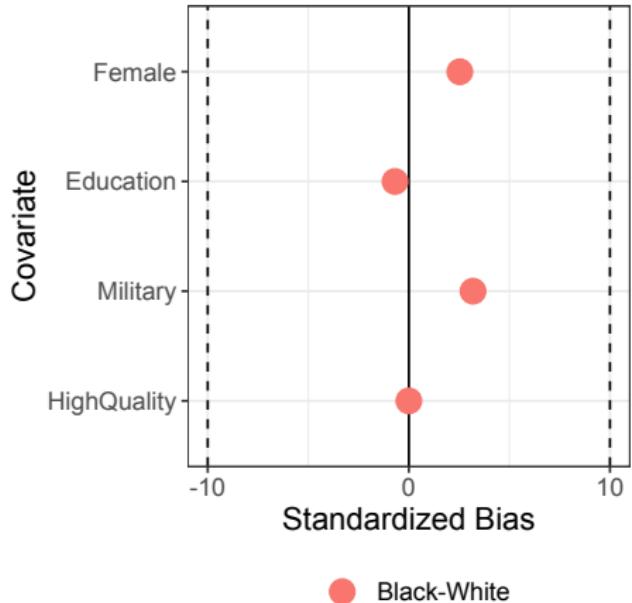
RCT の力

	白人の名前	黒人の名前
Female	76.42%	77.45%
High Quality	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、 χ^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

標準化差分について

$$\text{SB}_{T-C} = 100 \cdot \frac{\bar{X}_T - \bar{X}_C}{\sqrt{0.5 \cdot (s_T^2 + s_C^2)}} \text{ where } X \text{ is continuous,}$$
$$\text{SB}_{T-C} = 100 \cdot \frac{\bar{X}_T - \bar{X}_C}{\sqrt{0.5 \cdot (\bar{X}_T(1 - \bar{X}_T) + \bar{X}_C(1 - \bar{X}_C))}} \text{ where } X \text{ is binary.}$$

- \bar{X}_T : 処置群における X の平均値 (C : 統制群)
- s_T^2 : 処置群における X の分散 (C : 統制群)
- $|SB|$ が一定の基準を下回れば OK
 - 3 や 5(Caliendo and Kopeing 2008)、10(Normand et al. 2001; Austine and Mamdani 2005)、25(Ho et al. 2007) などの基準
- グループが 3 つ以上ならそれぞれのグループをペアに実行

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): $\mathbb{E}[\text{Call}|\text{Race} = \text{Black}] - \mathbb{E}[\text{Call}|\text{Race} = \text{White}] = \textcolor{red}{-0.032}$
- ATE = 0 の帰無仮説の検定
 - $t_{df=4711.7} = -4.117; p < 0.001; 95\%CI = [-0.047, -0.017]$
 - 統計的有意性 ($p < \alpha$) のみにこだわらないこと (Gelman and Stern 2006)

方法 2: 単回帰分析 (線形 or ロジスティックス/プロビット)

Table: 線形回帰分析 (LPM)

Coef.	Est.	S.E.
Intercept	0.064	0.006
Race:White	0.032	0.008

Table: ロジスティック回帰分析

Coef.	Est.	S.E.
Intercept	-2.675	0.083
Race:White	0.438	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

Coef.	Est.	S.E.
Intercept	0.057	0.007
Race:White	0.032	0.008
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

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	差分 (ATE)	t	p
	$(T = 0)$	$(T = 1)$			
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

因果効果の不均一性

方法2 性別と処置の交差項を投入した重回帰分析

	Est.	S.E.
β_0 : Intercept	0.611	0.091
β_1 : Treatment	0.951	0.131
β_2 : Female	-0.117	0.127
β_3 : Treatment \times Female	1.036	0.183

- 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

		Aさんの Outcome の例			
		Aさん		Aさん	
		統制	処置	統制	処置
Bさん	統制	0	10	Bさん	統制
	処置	15	20		処置

- **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) 隣の人が見てると私が見てるのが違いますが...?

実習

実習

これまで紹介した例題

① 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

② 処置効果の不均一性: Day1_Data2.csv

- ▶ 講師が作成した架空データ

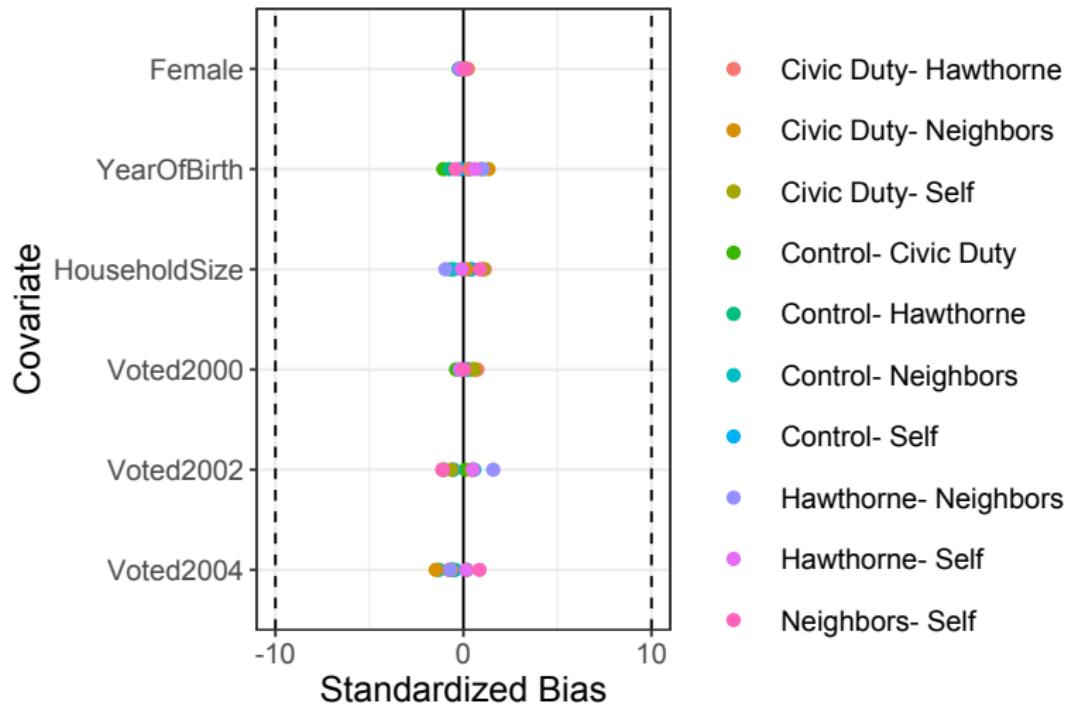
Lab Session 用のデータ

③ 社会的压力と投票参加: 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

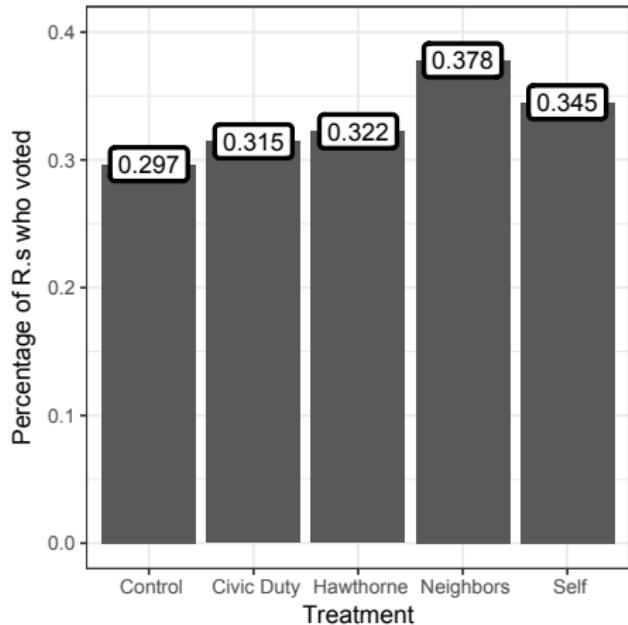
バランスチェック

処置が複数 (multiple treatment) であるため、組み合わせごとに標準化差分を計算

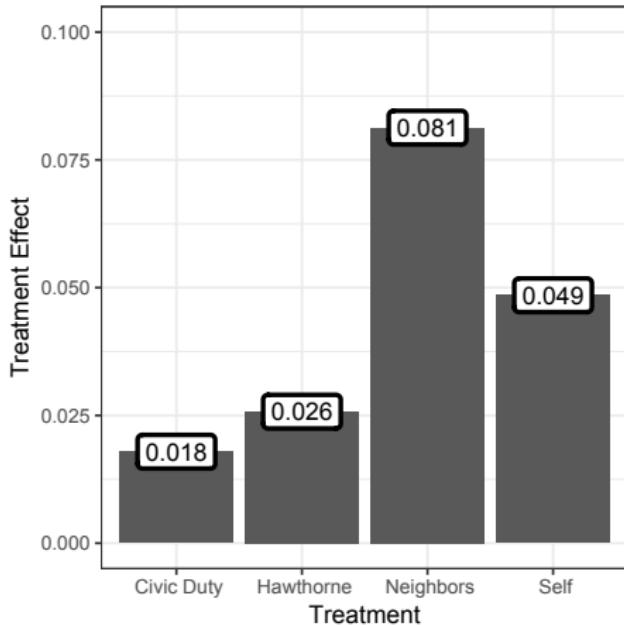


処置効果: 可視化

処置効果 (1)



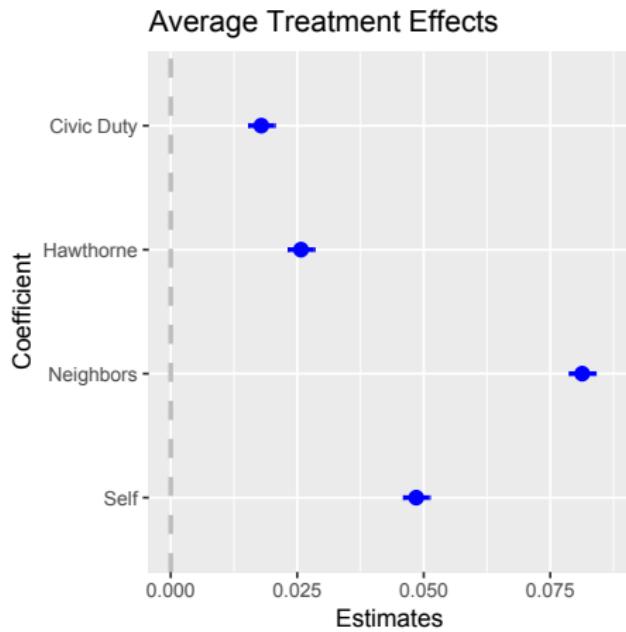
処置効果 (2)



処置効果: 回帰分析

- 4つの推定値は各処置 (Civic Duty, Hawthorne, ...) の平均処置効果 (ATE)

Coef.	Est.	S.E.
Intercept	0.297	0.001
Treatment		
Civic Duty	0.017	0.003
Hawthorne	0.026	0.003
Neighbors	0.081	0.003
Self	0.049	0.003
Obs.	344,084	
Adj. R^2	0.003	



分散分析

一元配置分散分析 (Oneway ANOVA) と多重比較

- 処置群間の比較に有効
- ペアごとの差の検定でなく、多重比較を行うこと

Table: Oneway ANOVA and Multiple Comparison

		(J)				
		Control	Civic Duty	Hawthorne	Neighbors	Self
(I) – (J)						
(I)	Control	0.000				
	Civic Duty	0.019 (< 0.001)	0.000			
	Hawthorne	0.026 (< 0.001)	0.008 (0.134)	0.000		
	Neighbors	0.081 (< 0.001)	0.064 (< 0.001)	0.056 (< 0.001)	0.000	
	Self	0.049 (< 0.001)	0.031 (< 0.001)	0.023 (< 0.001)	-0.033 (< 0.001)	0.000

$$F_{df=4} = 293; p < 0.001$$

Note: *p*-values in parentheses.