

プロキシを通して計量的“説明”を考える

樊怡舟* ・ 中尾走**

Yizhou FAN, Ran NAKAO

*広島大学高等教育研究開発センター 特任助教

fanyizhou921106@yahoo.co.jp

**愛媛大学次世代人材育成拠点 特定講師

nakao.ran.fq@ehime-u.ac.jp

自己紹介

樊 怡舟 (ハン イシュウ) Yizhou FAN

略歴

- 東華大学外国語学院日本語学科 学士(文学)
- 広島大学大学院教育学研究科博士課程前期修了 修士(高等教育学)
- 広島大学大学院教育学研究科博士課程後期修了 博士(教育学)
- 統計数理研究所 統計エキスパート育成プロジェクト研修 三期生
- 広島大学高等教育研究開発センター 特任助教

専門

- 高等教育論, 数理社会学, 計量社会学

研究関心

- 質的・量的研究法に関する方法論的検討
- 数理モデルに基づいた社会理論の検証
- 社会科学のための因果推論手法の開発と適用

自己紹介

中尾 走(なかおらん)

Ran NAKAO

略歴

- 愛媛大学工学部応用化学科卒業 学士(工学)
- 愛媛大学大学院教育学研究科学校教育専攻修了 修士(教育学)
- 日本学術振興会特別研究員(DC2)
- 広島大学大学院教育学研究科博士課程後期修了 博士(教育学)
- 広島大学高等教育研究開発センター 研究員
- 広島市立大学 大学評価・IRセンター 協力研究員
- 愛媛大学 次世代人材育成拠点 特定講師

専門

- 高等教育論, 数理社会学, 計量社会学

研究関心

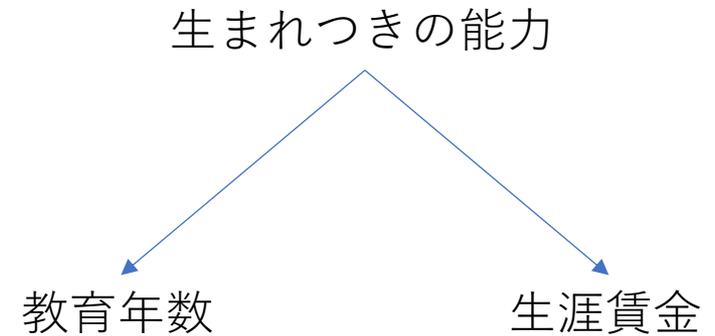
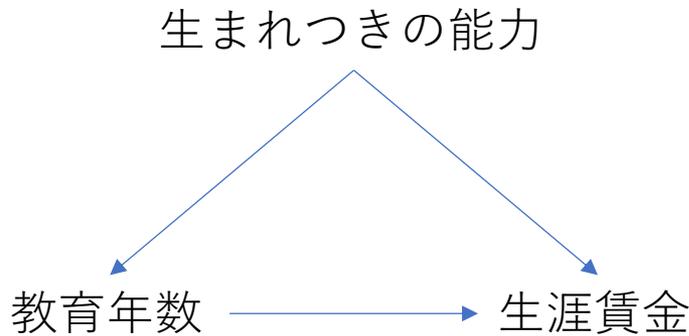
- プロキシを使った因果推論
- 測定誤差に関する経験データ取得
- (高等)教育に関する実証研究

社会科学の理論と実証における因果的「説明」の位置づけ

- 社会科学の理論研究すべての研究に因果の実証が不可欠で、慣習的にやられてきたもの（規定要因分析など）はほとんど価値を有しない（Samii, 2016） = Causal Empiricistの台頭・KKV論争
- 一方、狭い因果よりも広い「説明・解釈」の提出と検証が重要（Huber, 2013）
～介入操作が難しい課題・方法論的全体主義の課題など、因果推論を適用しにくい分野の衰退を招来しかねない
～**誘導型推定の旺盛により社会理論との断絶**が懸念される
- **因果的「説明」を意識した実証研究と社会理論との接合をどう考えるか？**
Deductive Nomological Model (DN Model)
Abductive Inference = Inference to the Best Explanation (IBE)
(Spirling, 2024)

相反する理論の識別問題

- 人的資本論 vs シグナリング理論



仮に生まれつきの能力など通常では観測できないものをデータとして手に入ったとしても、**仮設検定のパラダイムにおいて、シグナリング理論は立証され得ない...**

→ 「パスがない」ことを立証するのに、帰無仮説を積極的に受け入れる仕掛けが必要 = 反証主義を超えた仮定

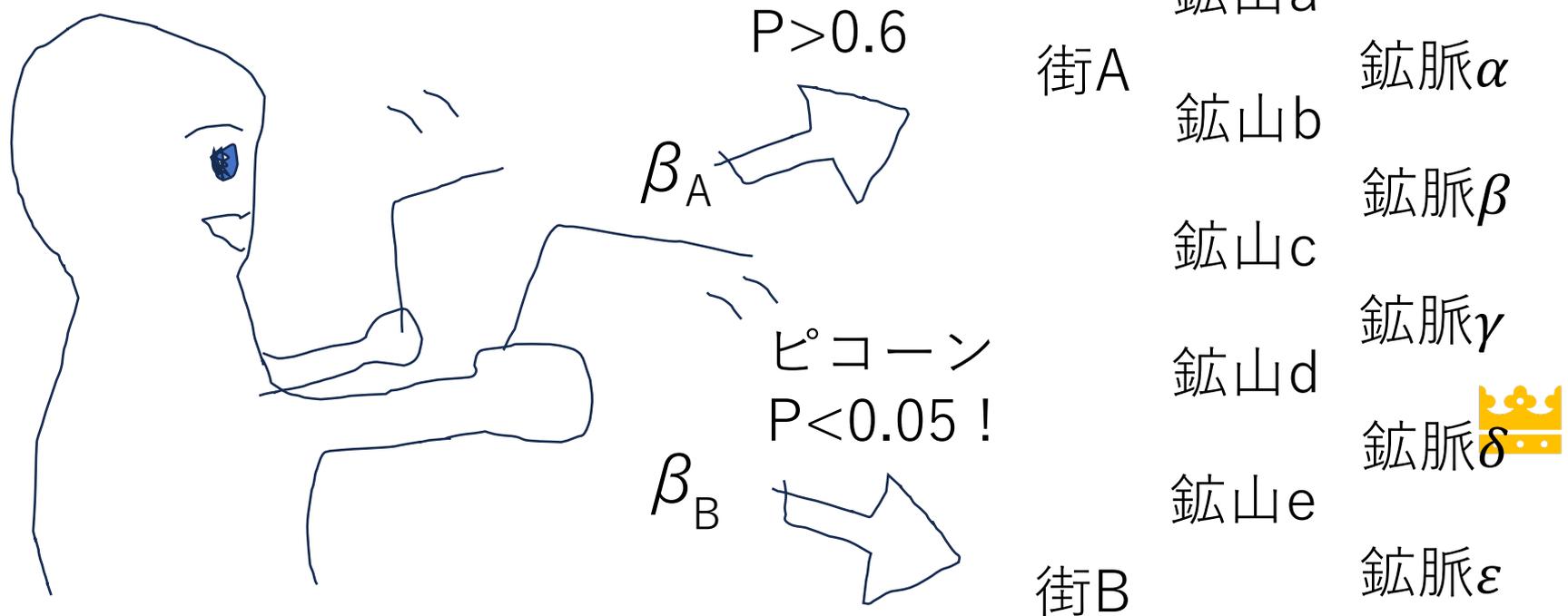
→ 理論の本質「教育が能力の成長につながるか」についての検証から遠ざかる

因果的「説明」の限界？

- 「説明」や「解釈」にもともと因果的意味合いが入っている。
DN model、Salmon model
～ 実証においてPearlの（局所）構造的因果モデルに収斂
- 因果推論の基礎づけ
＝「因果関係」の定常性・自律性（黒木，2017）
- 「each parent-child relationship in the network represents a **stable and autonomous physical mechanism**」(Pearl, 2009)
→因果推論＝データ生成過程を基礎づける物理的メカニズムの解明
- 社会科学の諸課題の最大な特徴
＝**抽象的な「構成概念」(Constructs)が中核**にある
～文化資本・学習能力・研究力・信頼感
- 「構成概念」から展開する社会理論は定常的・自律的なphysical mechanismとして捉えきれなのだろうか？**それを必要不可欠とするか？**
- リアリズムを根底にした因果的「説明」という視角が限界を迎えることに？
社会科学の課題に対応するようにどう拡張するか？

实用主義から見た準因果的「説明」

- 实用主義 (Pragmatism) 的視角
因果じゃなくても、研究課題への「理解」に資するものであれば、「説明」として成り立つ (Achinstein, 1983)
- 林先生のダウジングの例をお借りしますと...
方向づけ用の変数自身もダウジングの対象に
- * 「街A」という探索の指標と「街A」という存在とは違う意味をしている



構成概念と代理指標

構成概念

- * 抽象的な概念
- * 測定できないもの
- * 客観的定義できるかすら怪しい

理論変数

ギャップ

操作的定義

誤差項

代理指標

- * 定義・測定可能な指標
- * 抽象概念と関連している
- * 抽象概念の一側面を表している

実証変数

- 構成概念に基づいた社会理論に関する検討は、あくまで**プロキシ**を介して**間接的にアプローチ**するしかない
- 社会理論とデータ解析

理論モデルと実証モデル

理論変数

豊かさ → 教育熱心

実証変数

家庭年収 → 教育投資

- 「家庭年収→教育投資」の関係をデータに基づいて推論し、「豊かさ→教育熱心」について「説明」を実践する
- (精神的) 豊かさは教育熱心など文化資本に影響されることも想定できる ~ 「豊かさ→教育熱心」ではなく、「教育熱心→豊かさ」という理論モデルに→これを許すと、理論モデルが定常的ではなくなる →構成概念の定義を事前に定めないといけない
- **家庭年収を豊かさの指標として見立てているか、豊かさが家庭年収によって定義されるとみなすか？**
- 「豊かさ→家庭年収」ではなく、「家庭年収→豊かさ」に
- 科学的解釈が、単純なプロキシ間の条件付き関係性の記述に退化してしまうのではないか
- 社会科学の理論が追求しているのは、
Physical Mechanism or Cognitive Mechanism
Information Equivalence = Dafoe, Zhang & Caughey (2018)

プロキシへ注目

- プロキシを使うことが特徴な社会科学、その「説明」（プロキシを使用した「規定要因分析」？）はどのようなものか。
- 「規定要因分析」を、あえてプロキシを使用した因果推論の拡張として考える。→ **プロキシを使用した際の（不完全な）因果推論に対する感度分析という形で、観測されない「構成概念」間の関係性について理解を試みる。**
- この視点からができてきたことを考えていく。

- 中尾走, 樊怡舟, 村澤昌崇 (2021) 「交絡変数の測定誤差が合流点バイアスをもたらす影響について」『理論と方法』71: 34-52.
樊怡舟・中尾走, 2022, 「従属変数の測定誤差が統計的因果推論へ与える影響」『Advancement of Higher Education Research: RIHE Monograph Series』10: 1-19.
Yizhou FAN, Ran NAKAO (2023) 「One proxy for multiple unobserved confounders, what bias will be affected?」The 8th Japanese-German Symposium on Classification.
Ran NAKAO, Yizhou FAN, Kaixiang KANG (2024) 「Tradeoffs in an Endogenous selection bias due to missing and measurement error: In the case of confounder」Japanese Society for Quantitative Political Science (JSQPS) 2024 Winter Meeting Program.
中尾走, 樊怡舟, 康凱翔 (2024) 「Mediation Control Strategy」数理社会学会発表資料。

- 測定の精度ではなく、推論のバイアスに注目。
- 測定誤差を小さくすることが不偏な推定に繋がるか？

一例：社会的望ましさバイアス（SDB）

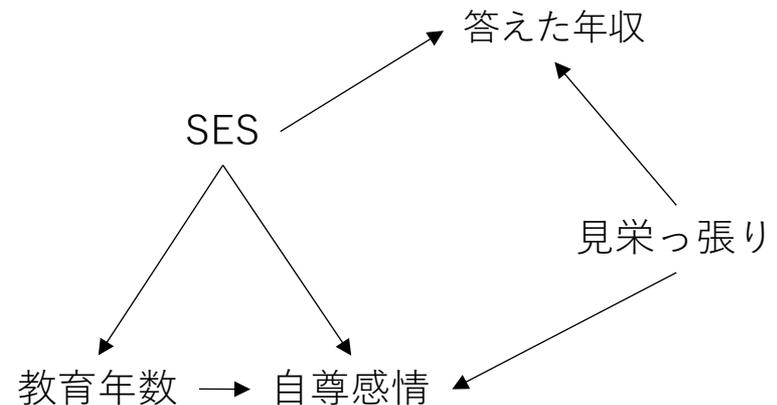
SESの代理指標として**家庭年収**を答えてもらうことが多い

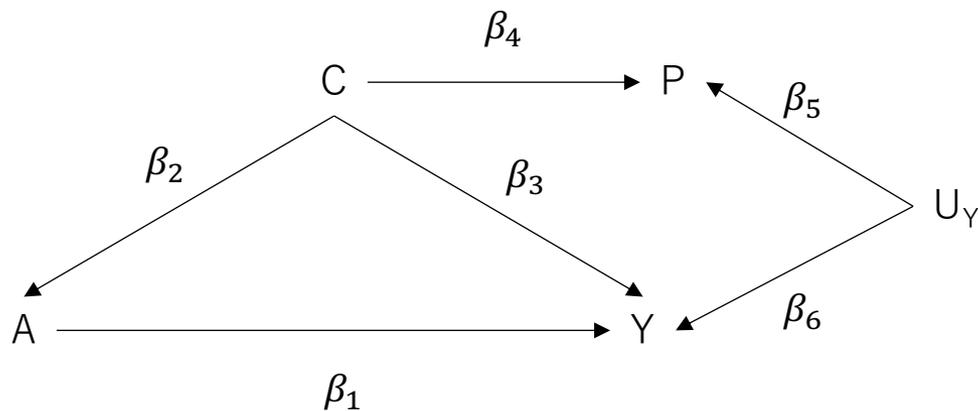
→「**答えた年収**」というプロキシ

見栄を張るなど一部水増して答える現象もみられる = 社会的望ましさバイアス

ただし、見栄張りは同時に従属変数である自尊心の一要因となりうる～Pの統制によって複雑な挙動を導入することになりそう

このような代理指標を統制することの是非を検討する = 逆に統制しないほうが推定値が真値に近いんじゃないの？





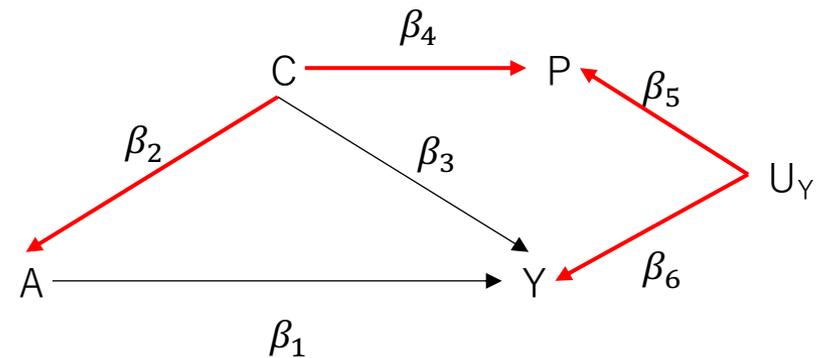
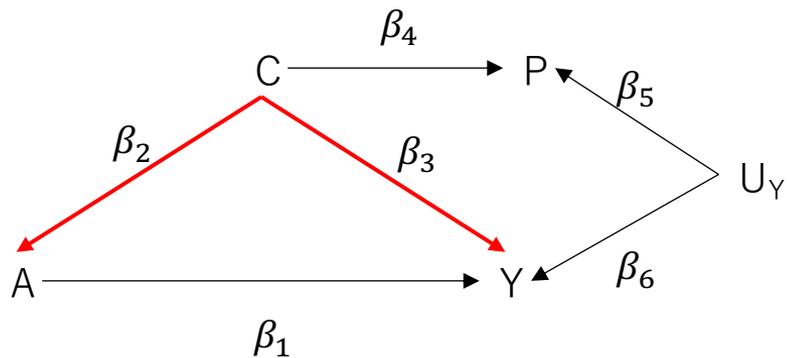
A、C、P、Yは標準化されていることとする。

$\rightarrow \text{var}(A) = \text{var}(C) = \text{var}(P) = \text{var}(Y) = 1$

β_1 を推定する際に、AとC（未観測）を線形回帰式に投入する必要がある

$$Y = b_1 A + b_2 C + a$$

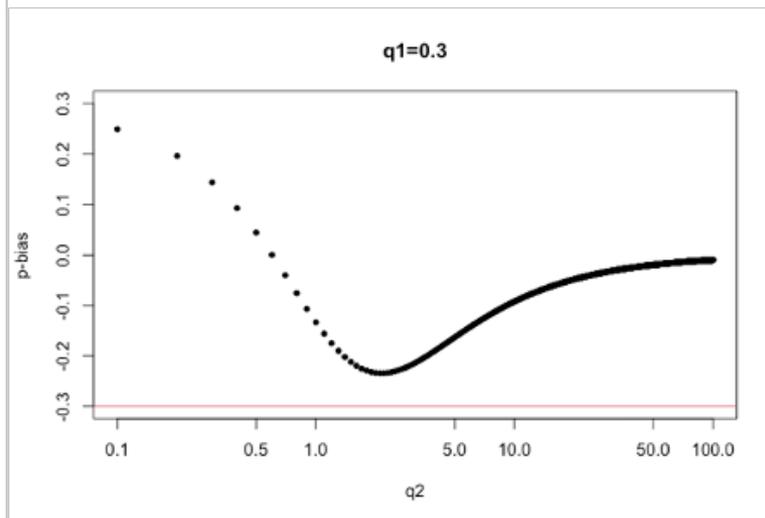
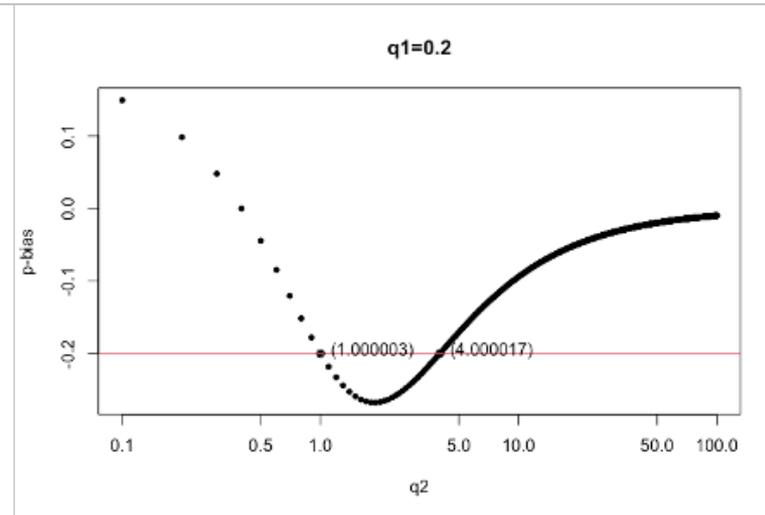
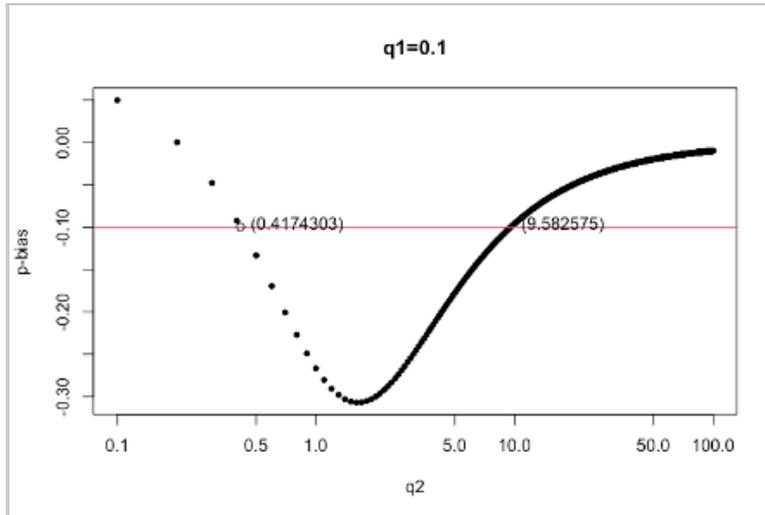
部分的統制と合流点バイアスのトレードオフ



Cの代理指標Pを代替的に回帰式に投入すると、Cが起因する内生性問題はある程度抑えられるが、 U_Y によるバイアスを新に導入することになる

→部分的統制 (partial control) と合流点バイアス (collider bias) とのジレンマ

測定精度上がると逆に推定の偏りが大きくなる可能性も

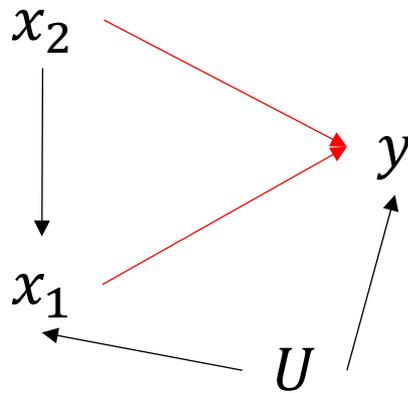


不完全な因果的「説明」の様式として、このように挙動する！
→誤差項を抑制すればいいというものではない。

この文脈の感度分析≠真値からのハズレに注目するわけではない
～不完全な「説明」から理論の標的となる構成概念下の「真値」を間接的に理解し得る。

バックドア基準を満たさないが…実用的？

- 全ての係数を正とする
- x_1 と y の間に未観測（未統制）の交絡変数がある場合, x_1 は合流点にもなる
- x_2 の効果を x_1 でブロックした方向と逆方向に戻す役割を担う
- この場合, 多少は過小推定されたとしても, x_2 の総合効果に近い推定値を赤線で得られているかもしれない



終わりに

- 因果的「説明」を目指さなくても、DAGの必要性・活用可能性が窺える。
- プロキシによる分析を意識した方が良い
- 観測されない・規範的定義すらされない「構成概念」に関して、感度分析を通した「理解」は可能かもしれない。

→因果推論を理解した上での「規定要因分析」が必要

そこで、PR!

11月5~7日 三日間丸ごと 因果推論研究集会 in愛媛大学

「統計的因果推論研究の最先端」

今日もご登壇いただいた林先生のほか、黒木学先生、佐藤俊哉先生、清水昌平先生、星野崇広先生、山本鉄平先生…各分野の統計的因果推論の第一人者ら計十五名から、最先端の因果推論の理論と実践をご説明くださいます。

参考文献

- Achinstein, Peter. 1983. The nature of explanation. Oxford University Press on Demand.
- Dafoe A, Zhang B, Caughey D. Information Equivalence in Survey Experiments. *Political Analysis*. 2018;26(4):399-416.
- Huber, John. 2013. “Is Theory Getting Lost in the ‘Identification Revolution’ ?” *The Money Cage* .
- 黒木学, 2017, 『構造的因果モデルの基礎』 共立出版.
- Pearl, J., 2009. Causality: Models, Reasoning, and Inference, the 2nd Edition. Cambridge University Press.
- Samii, Cyrus. 2016. “Causal empiricism in quantitative research.” *The Journal of Politics* 78(3):941–955.
- Spirling, Arthur & Stewart, Brandon M. "What Good is a Regression? Inference to the Best Explanation and the Practice of Political Science Research" *Journal of Politics*, conditionally accepted.