

# 労働経済学 2

川田恵介

## Table of contents

1	後期の内容	2
1.1	格差分析 . . . . .	2
1.2	集計指標を用いた分析 . . . . .	2
1.3	方針 . . . . .	3
1.4	予習 . . . . .	3
2	格差分析	3
2.1	データ分析工程の分解 . . . . .	3
2.2	分析工程 . . . . .	4
2.3	研究課題の設定 . . . . .	4
2.4	例 . . . . .	4
2.5	Estmand の定義 (識別) . . . . .	4
2.6	Estmand の定義 . . . . .	5
2.7	Estimand の推定 (推定) . . . . .	5
3	因果効果の識別	5
3.1	例. ゲーム時間規制 . . . . .	5
3.2	識別: 因果効果 . . . . .	5
3.3	因果についての枠組み: 実験アプローチ . . . . .	6
3.4	理想的な実験 (RCT) . . . . .	6
3.5	自然実験 . . . . .	6
3.6	Egami et al. (2024) の識別 . . . . .	6
3.7	実際 VS 理想の実験 . . . . .	7
3.8	Bad Control . . . . .	7
4	格差の識別	7
4.1	規範命題 . . . . .	7
4.2	例. 地域間格差 . . . . .	8
4.3	$X$ = 格差の定義 . . . . .	8

4.4	例: 世代間の格差移転 . . . . .	8
4.5	例: コホート内格差 . . . . .	8
4.6	例: コホート & 職業内格差 . . . . .	9
4.7	練習問題: コホート & 祖父母の学歴 . . . . .	9
4.8	格差の定義 . . . . .	9
4.9	他の例: Rose (2023) . . . . .	9
4.10	まとめ . . . . .	9
	Reference . . . . .	10

## 1 後期の内容

- 格差分析
- 労働市場の集計指標 + 経済理論を活用した分析
- とともに近年も方法的/概念的な議論が進む

### 1.1 格差分析

- 多くの家計にとって、労働所得が最大の所得源
  - 労働所得のばらつきは、社会全体の格差に大きな影響を与える
  - 経済全体の総所得に占める労働所得の割合は、労働所得への依存度が低い傾向にある”Top 1%”への所得集中度に大きな影響を与える
- 労働経済学のみならず、社会科学全般、公衆衛生等で大きな関心となっている

### 1.2 集計指標を用いた分析

- 経済学の伝統的な議論
  - 多様な経済主体間の相互作用に関心
  - 少数の指標に基づく、政策評価に懐疑的
- 古典的な因果推論: No interference (実験主体間で相互作用がない) および  $Y$  は観察できる を仮定しており、回答困難な局面も多い
  - 経済モデルの明示的な導入を議論 (他のアプローチ: [Wager \(2024\)](#) 11 章参照)
- マクロ経済学などと隣接

### 1.3 方針

- 講義資料は Github で共有: [レポジトリ](#) (講義期間中はブックマークしておいてください)
- 新規参加者もいる場合、前期の必要箇所を復習しながら議論
  - 前期の資料は 2024Spring から入手可能
- 講義の最後に、R の実習を行いながら復習

### 1.4 予習

- 格差分析について予習したい人には、以下を推奨
  - [Disparity analysis: A tale of two approaches](#)
- まだ読んだことない人は、以下のページを強く推奨
  - [Causation, Comparison, and Regression](#)

## 2 格差分析

- 規範的観点から望ましくない差 (差別/格差/不平等) について分析
- ポイント
  - Estimand の推定方法については因果効果の推定と多くを共有できるが、Estimand の設定方法が異なる
    - \* しばしば混同されてきた (Rose 2023)
  - 本日: 因果効果と格差分析の分析工程を整理 (復習)

### 2.1 データ分析工程の分解

- データ分析: 研究課題に、データを用いて回答する
- 具体的な分析工程に分解することが有益
  - 工程の見直しも行われてきた (Manski 1995; Lundberg, Johnson, and Stewart 2021)
    - \* 機械学習や Balancing weights、因果推論や格差の定義などの有効活用の大前提
- ポイント: 推定対象 (Estimand) の定義と推定の分離

## 2.2 分析工程

- 本講義では以下の工程を想定
  1. 応用したい意思決定問題に応じた研究課題の設定
  2. 推定対象の定義 (識別)
    - 因果推論、格差/不平等
  3. 推定方法の決定 (推定)
    1. Nuisance や Balance weight を OLS/最尤法/ベイズ/機械学習などを用いて推定
    2. Estimand をモーメント推定

## 2.3 研究課題の設定

- 比較研究:  $D$  間で見られる  $Y$  についての差を明らかにする
  - 例: 去年と今年 ( $D$ ) で平均賃金 ( $Y$ ) はどの程度変化したのか?
- 典型的な因果効果/格差分析は、特殊な比較研究
  - 因果効果の推定: 「 $D$  への介入によってもたらされたと差である」と解釈できる比較研究
  - 格差研究: 「 $D$  間での差は規範的に望ましくない」と解釈できる比較研究

## 2.4 例

- 因果:  $D =$  少なくとも原理的には、人為的な介入が可能な変数
  - 実証労働経済学の受講、大阪大学経済学研究科への進学
- 格差:  $D =$  何らかの集団 (Social Category)、介入が不可能でも OK、集団の定義に”科学的/客観的な根拠”がなくても OK
  - “性別”、“人種”

## 2.5 Estimand の定義 (識別)

- 推定対象の定義: 母集団上における差が、研究課題に回答可能なのか? を論じる
  - 極力”Nonparametric”に定義される
    - \* 比較研究:  $E[Y|D = 1, X] - E[Y|D = 0, X]$  の荷重平均

## 2.6 Estmand の定義

- $Y, D, X$  と分析事例の定義 = 研究課題に対応する推定対象の定義
- $Y/D$  の選択は、研究課題と比較的明確な関係性を持つ
  - 持たない場合は、解決は非常に難しい
- $X$  の選択については、多くの建設な議論が蓄積
  - **非推奨**：「 $Y$  に影響を与えそうな  $X$  を全て用いる」

## 2.7 Estimand の推定 (推定)

- 推定方法: 限られた事例数から、推定対象に近い値をどのように算出するか?
  - Balancing weight を機械学習/OLS/最尤法/ベイズで推定する
    - \* 限られた事例数に対応するために、適切な単純化を用いる
- 推定のため単純化と推定対象を定義を分離する

## 3 因果効果の識別

- $D$  を変化させる介入による  $Y$  の変化を推定したい
  - 差を推定したとして、介入の効果と解釈できるのか?

### 3.1 例. ゲーム時間規制

- 課題: ゲームの利用時間を規制すべきか?
  - ゲームが健康状態に与える因果的悪影響を与えるのか?
- 推定対象:  $E[Y|D = 1, X] - E[Y|D = 0, X]$ 
  - $Y$ : 精神的健康状態の指標
  - $D$ : “ゲーム” についての変数

### 3.2 識別: 因果効果

- 推定対象の定義:  $E[Y|D = 1, X] - E[Y|D = 0, X] = D$  への介入による変化であれば、平均的因果効果と解釈できる

- 複数の有力なアプローチ (潜在結果や構造的因果モデルなどがある)
  - 少なくとも、介入を想像することの重要性は共通

### 3.3 因果についての枠組み: 実験アプローチ

- 理想的な (思考) 実験 (Target Trial) に近づける
  - 本講義で採用
  - 他の枠組みの参照点としても機能

### 3.4 理想的な実験 (RCT)

- 無限の被験者が存在し、被験者間相互作用がない環境において、各被験者の  $D$  をランダムに決める
  - データから観察可能/不可能な背景属性と”無関係”に  $D$  は決定
  - $Y$  の分布に差があれば、 $D$  の違いによってもたらされたと解釈する”しかない”状況に持ち込める
  - 被験者数が有限であれば、背景属性は”偶然”偏るが、その偏りは信頼区間で評価できる

### 3.5 自然実験

- 社会の中で”勝手に起きた”実験的状况
  - 理想の実験をどうすれば近似できるか?
- $X$  として用いる変数: 理想の実験では、 $D$  間で違いが生じない変数
- 用いるべきではない変数: 理想の実験においても、 $D$  間で違いが生じる変数 (Bad controls/Post-treatment variables/Mediator)
  - 発展的議論 (VanderWeele 2019)

### 3.6 Egami et al. (2024) の識別

- コロナ下で生じたゲーム機への超過需要とそれに伴う”くじ引き”を活用
  - 購入希望者の中で、くじに当選した人しかゲーム機が買えなかった
- 分析対象事例: くじへの参加者
  - $D$  = 当選者
  - $X$  = 年齢、性別、子供の有無などの背景変数

- 背景属性を揃えて比較することで、理想の実験に近づける

### 3.7 実際 VS 理想の実験

- 理想: ゲーム機購入者を日本全体で完全ランダムに決めれば、背景属性はほぼ揃うはず
- 現実:
  - 参加者が限られているので、背景属性は偶然偏る
  - 居住地は一貫して異なる可能性
    - \* 需給が緩い地域の方がくじが当たりやすい
- 居住地も含めた背景属性を揃えて比較することで、理想の実験に近づける

### 3.8 Bad Control

- 古典的誤解:  $Y$  (健康状態) に影響を与えそうな全ての変数を揃える
  - 因果効果の経路を潰してしまい、効果を過大/過小に推定してしまう
- 例: ゲームのプレイ時間を揃えてしまうと、ゲーム機所有の因果効果は消失する可能性が高い
  - 「ゲーム機を持っておらずプレイ時間 0 VS ゲーム機を持っているのにプレイ時間 0」の比較になるため

## 4 格差の識別

- $D$  間で、本来的に許容すべきではない差を推定する
- 因果効果と異なり、多くの実証研究で用いられる枠組みを (川田は) 知らない
  - 因果推論とは、質的に異なる難しさ (規範命題) がある
  - 本スライドは、格差を因果的には定義しにくいという主張 (Rose 2023; Jackson and VanderWeele 2018) に従う

### 4.1 規範命題

- 格差 (差別、不平等)  $:=$  観察された差が格差であるかどうかは
  - 「である」(事実命題) ではなく、「べき」(規範命題)
- 客観的事実から、規範を導くことには、慎重であるべき (できない?)

- ヒュームの法則

## 4.2 例: 地域間格差

- 東京都で生まれたか、香川県で生まれたかで
  - 30歳時点での実質所得に差がある
    - \* 機会の不平等?
  - おいりの消費量に差がある
    - \* 食文化の多様性?

## 4.3 $X =$ 格差の定義

- $X =$  研究関心となる格差/差別/不平等を定義する変数
  - 価値判断へのコミット
  - ある  $X$  内で差がなければ、研究課題となっている格差/差別/不平等はないと定義する

## 4.4 例: 世代間の格差移転

- $$D = \{1(\text{両親とも高校卒以下}), 0(\text{どちらかは大学卒})\}$$
- 思考実験:  $E[Y|D = 1] - E[Y|D = 0] = 0$  であれば格差はない?
  - 社会全体で進学率が上昇する中で、年齢が若い世代は、親の学歴は高い傾向
    - \* 年功賃金の下で、賃金は低い傾向
    - \* 「年齢が高いと、両親の学歴が低く、賃金が高い」だけでは?

## 4.5 例: コホート内格差

- $X =$  生まれ年 をバランスさせた方が妥当?
- 思考実験:
 
$$E[Y|D = 1, 2000\text{年}] - E[Y|D = 0, 2000\text{年}] = 0$$
 であれば格差はない?
- 同じ年に生まれたのに生じる差 (Cohort 内格差) を格差と定義



#### 4.6 例: コホート & 職業内格差

- $X = \{\text{生まれ年, 仕事内容}\}$  の方が妥当?
  - 仕事内容まで揃えると、推定される格差は非常に小さくなるかもしれない
  - “職業分断” という格差を”伝導する”経路を潰している?

#### 4.7 練習問題: コホート & 祖父母の学歴

- $X = \{\text{生まれ年, 祖父母の学歴}\}$  の方が妥当?

#### 4.8 格差の定義

- 因果推論 (教育への介入が、次の世代に与える因果効果を明らかにしたい) のであれば、祖父母の学歴はコントロールすべき
  - ある世代へランダムに介入するのであれば、その上の世代の特徴はバランスしているはず
- “先祖代々” から続く格差 (教育に熱心な”家系”かどうか) に関心があり、親の学歴をその代理変数とするのであれば、コントロールすべきではない

#### 4.9 他の例: Rose (2023)

- 司法現場における格差研究: “人種” 間で量刑に差があるか?
- スピード違反が対象
  - 研究課題が「警察官が差別意識を持っているかどうか」であれば、違反速度はコントロールすべき
  - 「違法行為 (と認定される) を犯す可能性についての格差」であれば、コントロールすべきではない
    - \* ある Social category は、違法行為を犯さざるを得ない状況に追い込まれているかもしれない

#### 4.10 まとめ

- 因果推論でも格差推定でも、Estimand の定義をデータのみで行うのは困難
  - 現状、変数の定義を確認し、背景知識を用いながら、人間が判断するしかない
    - \* 因果と格差で、判断の根拠が大きく異なる
- どちらも、最終的には比較分析が必要なケースが多い

– OLS も有力な方法だが、それを補完/代替する優れた手法が多く開発されている

\* 因果と格差で、同じ手法が活用できる

## Reference

- Egami, Hiroyuki, Md Shafiur Rahman, Tsuyoshi Yamamoto, Chihiro Egami, and Takahisa Wakabayashi. 2024. “Causal Effect of Video Gaming on Mental Well-Being in Japan 2020–2022.” *Nature Human Behaviour*, 1–14.
- Jackson, John W, and Tyler J VanderWeele. 2018. “Decomposition Analysis to Identify Intervention Targets for Reducing Disparities.” *Epidemiology* 29 (6): 825.
- Lundberg, Ian, Rebecca Johnson, and Brandon M Stewart. 2021. “What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory.” *American Sociological Review* 86 (3): 532–65.
- Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Harvard University Press.
- Rose, Evan K. 2023. “A Constructivist Perspective on Empirical Discrimination Research.” *Journal of Economic Literature* 61 (3): 906–23.
- VanderWeele, Tyler J. 2019. “Principles of Confounder Selection.” *European Journal of Epidemiology* 34: 211–19.
- Wager, Stefan. 2024. “Causal Inference: A Statistical Learning Approach.” preparation.