

大阪府×大阪公立大学 共催セミナー

知りたいことを知るためのアンケート 調査活用術

2024/11/08

黒田 雄太 (大阪公立大学)

自己紹介

• 略歴

- 1989年7月 山形県生まれ
- 2019年3月 博士取得（経済学・東北大学）
- 2019年4月 東北大学 大学院経済学研究科 助教
- 2020年10月 大阪市立大学 大学院経済学研究科 准教授
- 2022年4月 大阪公立大学 大学院経済学研究科 准教授

• 研究領域

- 教育経済学, 公共経済学に関連する実証分析（応用計量経済学）

• 連絡先

- kuroyu0725@gmail.com

自己紹介

• 最近（直近3年）の主要論文

- Kuroda, Y., & Sugasawa, T. (2023). The value of scattered greenery in urban areas: A hedonic analysis in Japan, *Environmental and Resource Economics*, 85(2), 523–586.
- Kuroda, Y. (2022). The effect of pollen exposure on consumption behaviors: Evidence from home scanner data, *Resource and Energy Economics*, 67, 101282.
- Kuroda, Y. (2022). What does the disclosure of school quality information bring? The effect through the housing market. *Journal of Regional Science*, 62(1), 125–149.

• 最近（直近3年）の科研費

- 若手研究「非選抜的な大学の収益に関する実証研究」2024年4月–2028年3月（代表）
- 基盤研究B「学校間競争と子どもの人的資本形成：教育施策の自治体間差異を利用した分析」2023年4月–2028年3月（分担）
- 挑戦的研究（萌芽）「オルタナティブデータを利用した学校教員の長時間労働の経済分析」2023年6月–2026年3月（分担）
- 若手研究「教育の地域格差に関する実証研究：公立・私立学校の役割の違いに注目した分析」2021年4月–2024年3月（代表）

今日の内容

1. 社会調査とはなにか
2. 社会調査の作成・実施
 - サンプルサイズ
 - 設問の作成
3. 調査結果の可視化・分析
4. 社会調査の活用・注意点
5. まとめ

今日やること/やらないこと

• やること

- 社会調査の実施に関する注意事項・TIPS集
- 統計処理の基礎的な考え方
- データ分析でありがちな問題点

• やらないこと

- 実施上の具体的な手順
- 言葉遣いなどの倫理・道徳的な配慮
- 統計的・数学的な詳細
- 統計ソフトの使い方やプログラミングのコードなど

社会調査とはなにか

- 質問票を用いてなんらかのデータを取得すること
 - いわゆるアンケート (questionnaire)
- 以下のようなプロセスで進展することが多い
 - 企画
 - 設計
 - 実施
 - 集計
 - 分析
 - 公表

なぜ社会調査をするのか

- **なんらかの情報を集めて「使いたい」**
 - 集めただけでは意味がない
 - 目的のない「やっただけ調査/分析」は世の中に多い
 - とりあえず「見える化」しろ
 - 最新の統計手法を使ってみろ
- **社会調査やデータ分析は「目的」が重要**
 - 目的が明らかになれば必然的にとるべき手段が決まる
- **なんらかの意思決定に結び付けることが大切（経済学では）**
 - 結果の一般化・外的妥当性を考慮する

社会調査の目的

- **目的（ゴール）を明確にしておくことが最重要**
- **社会調査にはコストがかかる**
 - なにをどこまでやるのか決めないと際限がない
- **入手したデータを後から追加・変更することが困難**
 - 調査後に「変数が足りない」となってもどうしようもない
- **その後の利用目的によってとるべき手段が異なる**
 - 取得したデータを何に使用するのか（研究/政策立案/内部報告）
 - 結果の公表や研究利用には事前に承諾が必要
 - 長期的な視野も重要

社会調査の限界

- **欲しいデータを得るためには適切な手段が必要**
 - 様々な懸念事項/落とし穴が存在する
- **万全を期してもどうしてもないことは多い**
 - アンケートは万能ではない
 - 他のデータや統計調査と補完的に使用する必要がある
 - そもそも測れないものは測れない
- **悪いデータからは悪い分析しかできない**
 - Garbage in, garbage out
 - データの取得段階で適切な処理をする必要がある

結局どうすればいいのか

- **まずは調査の目的を明確にする**
 - 目的によって適切な手法が異なる
 - 母集団の特徴を把握したい → 量的な情報が重要
 - ニーズや問題を洗い出したい → 質的な情報が重要
- **目的に応じた手法を選択する**
 - 統計や社会調査に関する知識が重要
- **それぞれの手法の利点/不利点を理解する**
 - どのような調査方法でもバイアスを排除することは難しい
 - バイアスの存在/傾向を認知し、解釈に反映させる

サンプリングの基本

- **社会調査は母集団から標本を抽出する作業**
 - 抽出方法は多数ある（単純無作為抽出法や層化多段抽出法など）
- **母集団から無作為にサンプルを抽出することが重要**
 - 偏った標本抽出はバイアスの原因
- **現実では高確率で偏りが発生する**
 - 調査票をランダムに送付しても回収率には差が生じる
- **統計・計量分析はバイアスとの戦い**

サンプリングにおける代表的なバイアス

• 選択バイアス

- サンプルが偏っている/回答しやすい人だけが回答している
 - インターネット調査だと若年者に偏る
 - 文字数が多いと忙しい人や高齢者が無視する

• 確証バイアス

- 自説に都合のいい解答だけを集計する・作成する

• 交絡バイアス

- 交絡因子が影響しているのを見落としている

• バイアスを除去するのは非常に大変 (RCTなどが必要)

サンプルサイズの問題

- **どのくらいのサンプルサイズが必要なのか**
 - 個人的に最もよく聞かれる質問
- **決め方には複数の観点がある**
 - 統計的視点
 - 標準誤差が小さくなるように調べる
 - 実務的視点
 - 「分析単位ごとに50サンプル欲しい」と言ったことがあり得る
 - (予算的視点)
 - サンプルサイズや設問数に上限がありえる

統計的観点

- 許容できる誤差の大きさから統計的に計算することは可能

$$n \geq \frac{N}{\left(\frac{\varepsilon}{K(\alpha)}\right)^2 \frac{N-1}{P(1-P)} + 1}$$

- α : 母集団特性値の推定を誤る確率 (5%なら $K(\alpha) = 1.96$)
- ε : 標本比率につける幅 (\pm)
- N : 母集団の大きさ, n : 必要なサンプルサイズ, P : 母比率

統計的観点

- 人口10万人の都市におけるX法案に賛成の人の割合を信頼区間95%で誤差の幅を±5%で推定したい場合（賛成率は50%と仮定）

$$n \geq \frac{100000}{\left(\frac{0.05}{1.96}\right)^2 \frac{100000-1}{0.5(1-0.5)} + 1} = 383$$

- **回収率が100%ならサンプルサイズは最低でも383必要**
 - 母集団の大きさ N は必要なサンプルサイズ n にあまり影響しない
 - 誤差の幅を小さくしたい場合はより多くのサンプルサイズが必要
 - 回収率を60%と想定した場合は $383/0.6 = 638$ 必要になる

実務的観点

- **分析目的によって収集すべきサンプルサイズは異なる**
- **均等割付**
 - 30代100人, 40代100人, 50代100人のように各世代から100人ずつ集計
 - グループ間の比較がやりやすくなる
 - 母集団の平均に関する推測が難しくなる (若年層の影響が大きくなる)
- **母集団構成比に合わせた割付**
 - 各世代を人口で重みづけし, 若年層を少なく, 高齢者を多く集計
 - 母集団の傾向が見やすくなる
 - グループ間の比較が難しくなる場合がある
 - 若年層同士の比較をしたいのにサンプルサイズが小さくて推定できないなど

質問の仕方

- **アンケートは質問者と回答者の対話**
 - 細心の注意を払って作成する必要がある
 - 相手に期待しすぎると失敗する
- **一方通行のコミュニケーションなので齟齬が生じやすい**
 - 基本的に回答者の善意や熱意に委ねられている
 - 回答が義務ではない場合はさらに深刻
 - 少しでも面倒/わかりにくいことがあると適当に回答する/回答をやめる
- **徹底的に事前調査/準備をして, 何度もプレテストする**

設問作成でありがちな失敗

- **回答者に過度な負担をかける**
 - 専門用語の多用, 長すぎる文章, 理解しにくい表現など
 - 記述式の回答を多用 (これは集計の負担にもなる)
 - 多すぎる設問数 (回答率と設問数はトレードオフ)
- **分析を見据えていない選択肢の設定**
 - 選択肢に抜けや漏れがある, 選択肢が多すぎる
 - 得られた回答が解釈できない質問 (曖昧な定義の質問など)
- **質問者が聞きたい順番/思いついた順番に質問する**
 - 回答の順番は非常に重要

客観的・主観的な情報

• 客観的実態

- 年齢や性別, 出身地など
- 直接的に尋ねても大きな問題は生じにくい
- 人によって認識に幅がある変数については要注意 (職業など)

• 主観的行動・主観的意識

- 日々の運動習慣, 食生活, 好きなものなど
- どのように尋ねるかによって回答が大きく異なる
 - 客観的統計資料と分布等を比較して妥当性を検証することも可能
- あえて主観であることを利用した分析も可能
 - 意識と行動の間にどのくらいの差があるかの検証など

定量的・定性的な情報

- **定量的な情報/質問**

- 身長や体重, 主観的健康度など
- 「あなたの主観的な健康度は1~10で表現するといくらですか？」

- **定性的な情報/質問**

- 「健康上の問題点がありますか？」 「~をする理由はなんですか？」

- **独自の社会調査では定性的情報が取得できるのが利点**

- 定量的情報は他の主要統計で補足できる場合が多い
- 潜在的な認識/意識の違いを把握できるがバイアスが含まれる
- 両者を組み合わせて使うことが重要
 - 定性調査で問題点を洗い出し, その問題について定量調査するなど

設問設定や回答の落とし穴

- **設問設定には対処すべき様々な懸念事項がある**
 - 実態と意識の乖離
 - 善意の嘘・タイトルバイアス
 - 無意識の誘導・同意バイアス
 - 言葉遣い（ワーディング）
 - ダブルバーレル質問
 - 機能不全な選択肢
 - 中心化傾向や比較
- **回答を集計/解釈する場合にも注意点多い**
 - 回答の正確性やバイアスの考慮
 - 無回答の罨

実態と意識の乖離

- **人間は必ずしも自分の意識通りに行動するわけではない**
 - ダイエットしたい人が8割 → 実際にする人は3割
 - 「〇〇が欲しい」は本当に〇〇の需要を反映しているのか？
→ 完全に的外れな施策を行ってしまうこともある
- **実態と意識の乖離を認識して調査に織り込むことが重要**
 - 定性的調査では需要があるのに定量的には反映されていない → なぜか？
 - その原因をさらに掘り下げることが有効な分析になりえる
→ 時間が無い, お金が無い, 心理的障壁があるなど
- **目的に沿った設問設定が必要**

善意の嘘・タイトルバイアス

- 回答者が本心ではない回答をする場合がある
- 質問者に忖度した回答
 - タイトルや設問から調査や実験の意図を推察する可能性がある
 - 「〇〇を問題視している調査だから〇〇に批判的な回答をしよう」
- 社会的に望ましいと思われる回答（社会適応性バイアス）
 - 本当は興味がないのに「環境問題は重要だ」と回答するなど
- 自分の見られ方を意識した回答（匿名であっても）
 - 所得を盛って記入する（低所得の場合は回答しない）
 - 趣味で行った行動を「社会貢献のため」という

無意識の誘導・同意バイアス

- **質問の仕方によって回答が誘導されうる**
 - オープンクエスチョン「最近の悩みを教えてください（選んで）ください」
 - クローズドクエスチョン「〇〇で悩んでいますか？」
 - ➔ オープンでは上がらないのにクローズドだと同意率が高いこともある
- **タイトルや設問順番の影響は大きい**
 - 「タイトル: 〇〇が心理的ストレスに与える影響に関する調査」
 - 「心理的なストレスを感じていますか？」 ➔ YESと答える人は多い
 - 「それは〇〇のせいですか？」 ➔ 流れでYESと答えてしまいがち
- **タイトルの有無や設問順番, 質問形式は熟考する必要がある**
 - 順序バイアス, アンカリング, キャリーオーバー効果などを考慮

無意識の誘導・同意バイアス

- **過剰（不必要）な説明**

- 「近年〇〇が社会問題化しており、～を調べる必要がある」
- 「あなたは〇〇について普段から意識していますか？」 → YESを誘導
- 「△△は～等の様々な健康被害をもたらすことが知られているが、あなたは△△についてどう思いますか」 → 否定的な回答を誘導

- **威光暗示効果**

- 「東京大学のA教授の調査では××だとわかっている」
- 「これについてあなたはどう思うか」 → ××という回答を誘導

- **誘導的な設問で得られた回答は信頼性が低い**

言葉遣い（ワーディング）

- **回答者によって解釈が異なる言葉を使用するのは危険**
 - 「身近な友達は何人いますか」 → 友達の定義は人それぞれ
 - 「ひと月に1回以上対面で会う友達は何人いますか」 → やや客観的
- **主観的な指標にも注意が必要**
 - 「主観的な幸福度は1～10のうち何点ですか」 → 人によって基準が違う
 - ただし他の統計指標や研究結果を参照して利用できる場合もある
 - 客観的な指標（自殺率や精神疾患の有無など）との相関が強いなど
 - 主観的な指標は使い分けが重要

ダブルバーレル質問

- 複数の要素を同時に含む質問をするべきではない
- 「飲酒喫煙はやめるべきだと思うか」
 - 「飲酒」と「喫煙」は必ずしも同時に回答できるものではない
 - 「飲酒には賛成」で「喫煙には反対」の人は答えようがない
- 「喫煙は健康に悪いのでやめるべきだと思うか」
 - 「喫煙は健康に悪いかどうか」と「喫煙をやめるべきか」は別の問題
 - それぞれに対して個別の回答が必要になる

機能不全な選択肢

- **選択式の設問の場合, 選択肢の設定は非常に重要**
- **よくない選択肢の例**
 - 相互排他的ではない選択肢 (一つを選ぶ場合)
 - 「りんご, みかん, ぶどうのうち, 好きな食べ物はどれですか」
 - 網羅的ではない選択肢
 - 「あなたの最終学歴は以下のうちどれですか」
 - 「中卒, 高卒, 大卒, 大学院卒」 → 高専や専門学校等が含まれていない
 - 多すぎる選択肢
 - 回答者が適当に回答する/回答を間違える可能性が高まる
 - 各選択肢に含まれる人が少なすぎて分析できなくなる
 - 「該当するもののうち上位n個を選んでください」系の選択肢
 - 必ずしも悪くはないが分析や解釈が非常に困難

中心化傾向や比較

- **n段階評価では極端な選択肢を避ける傾向にある（中心化傾向）**
 - 1（非常に悪い）～ 5（非常に良い）の選択肢では2～4を選びがち
 - 回答を恣意的に誘導/解釈する事例がある
 - 「〇〇には賛成か（5段階）」 → 「非常に賛成（15%）」 「まあ賛成（60%）」
 - 「〇〇を強く推している人は15%しかいない」 「〇〇の支持率は3/4もある」
- **他調査や他国との比較を検討している場合は特に注意**
 - スコア付の感覚が大きく異なる場合もある
 - 例えば10点満点の評価の場合
 - A国「基本は満点の10点で、何か問題があれば7～9をつける」
 - B国「基本は真ん中の5～6点で、良ければ7～9、悪ければ2～4、1と10はほぼ無し」
 - 平均だけではなく分布を確認することが重要

回答の正確性やバイアスの考慮

- **不真面目な回答や明らかなミス of 除外/修正が必要**
 - Web調査の場合, 回答時間はひとつの判断基準になる
 - ずっと同じ選択肢だけを選んでいる回答なども除外候補
 - 他の統計と接続/比較できる客観的指標を聞いておく
 - 誤入力や異常な分布の確認に使用できる
 - 注意力テストも有効
 - 序盤と終盤に同じ質問をして同じ回答をするかどうか
 - 「これには2を回答してください」などの質問にきちんと答えているか
 - あまりしつこいと回答者が脱落するおそれもあるので注意
- **設問や選択肢の順番によるバイアスの除外も大切**
 - 設問をランダムに並び替えるのも有効
 - 選択肢のランダム化はミスとのトレードオフ

無回答の罣

- **無回答を安易に除外することは危険**
 - (潜在的に) 他の選択肢の回答者がランダムに無回答なら問題はない
 - 特定の選択肢の回答者ほど無回答の傾向がある場合はバイアスがかかる
 - 年収の低い人ほど無回答を選択 → 観察可能な年収が上方に偏る
- **分布の確認は統計調査において非常に重要**
 - 他の悉皆調査と見比べて同様の分布になっているか確認

分析の前に

- **データの確認**

- いきなり分析を始めるのは危険
- プロット（ヒストグラムや散布図）や要約統計量の確認が大切

- **データのクリーニング**

- 異常値/外れ値の確認/修正
 - 最終的には目視判断になる
- 変数の集約/作成
 - 連続変数を二値変数にするなど
 - 恣意的な分析にならないように複数の集計を行う
- 自由回答の処理
 - 表記ゆれの修正
 - 変数の作成やテキストマイニングによる変数の検討

データの確認

- **図示と要約統計量のどちらも重要**
 - 図示（ヒストグラムや散布図など）
 - 視覚的にデータを捉えられるが、解釈や比較が困難
 - 要約統計量（平均値や相関係数など）
 - 理解や比較が容易だが、多くの情報が欠落している
- **データの分布によって統計量の意味は大きく異なる**
 - 平均点が5
 - 5を中心の正規分布なのか、1と10が多い凹のような分布なのか
 - 相関係数が0
 - 相関係数は線形関係しかとらえられない → U字の関係があるかもしれない

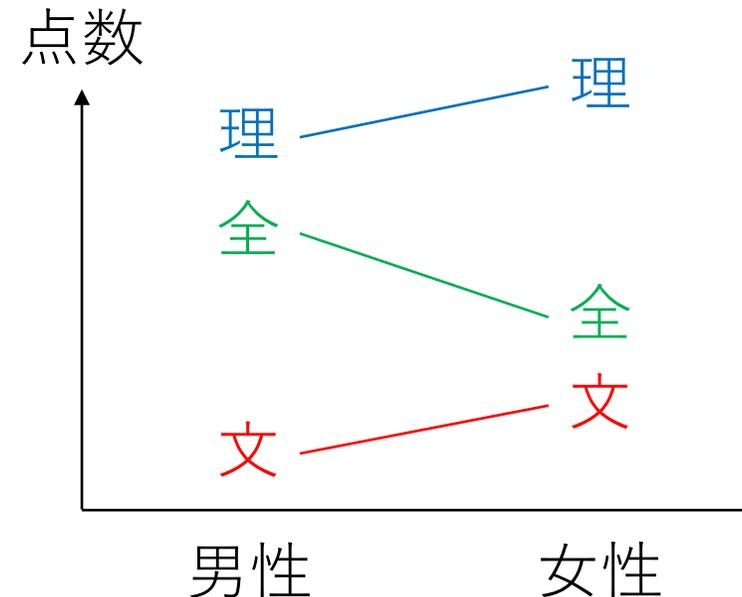
シンプソンのパラドックス

- サンプル全体の相関とサブサンプルの相関が逆になることがある

	理系	文系	全体	合格者
男性	80点 (90人)	60点 (10人)	78点 (100人)	70人
女性	90点 (10人)	70点 (90人)	72点 (100人)	60人

- 最終的な目的に応じて解釈が変わる

- 男女の平均的な学力差を見たい
→ 全体平均が重要
- 各分野における男女差を見たい
→ サブサンプル平均が重要



選抜効果

- **全体における相関関係が選抜された集団でみられないことがある**
 - 特定の変数に基づいて回答者を選定した場合に注意
 - ➔ その変数と他変数の相関関係がみられなくなる可能性がある
 - 例) 入試の成績と大学のGPAに相関関係がなかった
 - ➔ 入試の学力とGPAは無関係か？
 - ➔ 入試によって同程度の学力の生徒が選抜されている (差がない)
- **調査対象はどういう集団なのかに注意**
 - 想定される母集団はどういうものか
 - 調査対象の選択にはどのような条件があるのか

分析の方向性

- **1 変数についてプロット**
 - これだけからもわかることは多い
- **2 変数のクロス集計/有意差の検定**
 - もっとも簡単で、統計的な知識がなくても理解できる
 - 交絡因子によるバイアスが含まれる
 - ➔ 実務ではよく見るが、(経済学の) 研究では予備分析レベル
 - 直接的になんらかのエビデンスにするには弱い
- **3 変数以上の分析**
 - 多変量解析 (重回帰分析など) が必要
 - 3変数以上になると人間には直感的理解が難しくなる

統計分析の注意点

- **統計分析は万能ではない**

- データや指標は現実の一側面を捉えているにすぎない

- **調査には必ず誤差がある**

- 一度の調査結果のみで物事を判断するのは危険
 - 悪天候の日に調査をすると幸福度が下がる

- **「指標」は不変でも普遍でもない**

- 「精神疾患の患者が増えた」
 - ➔ 本当にその症状の人が増えたのか, 精神疾患の定義が変化したのか
- 「日本の一人当たり弁護士の数はい少ない」
 - ➔ 弁護士の定義は国によって違う (行政書士が弁護士に含まれる国も)

統計分析の注意点

- 「指標の変化」の解釈は難しい
 - 「犯罪件数が増加した」
 - ➔ 犯罪件数が増えたのか、警察の努力で犯罪の発見数が増えたのか
- 「測りすぎ」による問題
 - 測れるものだけを測ってしまう
 - 「頭の良さ」は測れないので「学力テストの成績」や「知能指数」で代理する
 - 測ったものだけが評価の対象になる
 - 本質的には「頭の良さ」を改善すべきだが「テストの成績」のみが評価項目になる
 - 測れないものや測りにくいもの（性格など）がおざなりになる可能性がある
- 「データからは何も言えない」という答えはある

まとめ

- **社会調査（アンケート）は量的/質的情報の取得に有益**
 - 調査にはコストがかかるので実施前に慎重な検討が必要
 - 調査の目的を意識することがなにより大切
- **統計的/実務的に様々な懸念事項がある**
 - 適切に対処しないと使えないデータが集まる
 - 様々なバイアスや心理学的傾向を知っておくことが重要
- **目的とデータに応じて適切な分析手法を選択する**
 - まずはプロットしてデータを確認するのが良い

社会調査のコツ（のようなもの）

- **まずは徹底的に関連資料や先行研究を調査する**
 - 「コストが高い」「やりなおしがきかない」ということを肝に銘じる
 - よい調査/研究を真似することが上達の近道
 - 「なぜこの設問があるのか/ないのか」の意味を考える
 - 先行研究の課題/欠点を補完するような調査には価値が高い
- **経験者を利用する**
 - 経験からしか学べないこと（暗黙知）は多い
 - 社会調査経験がある人の力を借りるのは非常に有益
- **回答者を信じすぎない/回答者に期待しすぎない**