

2024年度 統計データ分析コンペティション

## 審査員奨励賞 [大学生・一般の部]

データ駆動型因果探索による投票率変動要因の解明  
—全 1639 市区町村の投票率データを使用して—

関屋 百々花（一橋大学社会学部社会学科）

上原 颯馬、橋口 裕平、馬場 健生

（一橋大学商学部経営学科）

# データ駆動型因果探索による投票率変動要因の解明 —全 1639 市区町村の投票率データを使用して—

関屋百々花\*1・上原颯馬\*2・橋口裕平\*2・馬場健生\*2

\*1: 一橋大学社会学部社会学科

\*2: 一橋大学商学部経営学科

## 1. 研究のテーマと目的

### 1.1 背景と研究の目的

日本の国政選挙の投票率は、令和 3 年 10 月に行われた第 49 回衆議院議員総選挙で 55.93%、令和 4 年 7 月に行われた第 26 回参議院議員通常選挙では 52.05%となっている。衆議院選挙では昭和 33 年の 76.99%、参議院選挙では昭和 55 年の 74.54%を最後に国政選挙の投票率は減少傾向である。国政選挙の投票率の低下は国民の意思が適切に政治につながらず、政治が持つ代表性の低下をもたらす。投票率の低下に起因する代表性の低下は、日本という民主主義国家の根幹を揺るがす問題であり、早急な是正が求められる喫緊の課題であると言える。本研究は投票率の変化の要因となりうる変数を特定し分析することで、投票率を上げるための取り組みに貢献し、延いては日本の国政選挙の投票率向上に資するものである。

### 1.2 先行研究

投票率の要因に関する研究は十分とは言えないものの、いくつか行われている。例えば、柿本ら (2004) <sup>(1)</sup> では、投票率に対する期待効用モデルである Riker モデルをもとに、投票行動に影響を与える変数を明らかにしており、接戦度、一人当たり行政投資額、若者率が強く影響を与え、一票の格差、完全失業率も若干の影響を与えていることを明らかにしている。この事実をもとに、一票の格差是正、情報公開の推進、教育の充実を提言している。

## 2. 研究の方法と手順

### 2.1 分析対象変数の選定

先行する柿元ら (2004) <sup>(1)</sup> の研究において、Ricker モデルに基づき選定された変数を参考に、投票率を除き 10 の変数を選定した。柿元らの研究と同様に、完全失業率、第 3 次産業就業者割合を用いるとともに、年齢、職業の投票行動に対する影響をより詳細に探索するため、15 歳未満人口、15~64 歳人口、65 歳以上人口、第 1 次産業就業者割合、第 2 次産業就業者割合を追加した。また、新たな変数として、歳出に対する民生費割合、転入者、転出者の総人口に対する割合を追加した。

### 2.2 投票率の計算

加えて、2022 年度の第 26 回参議院議員選挙の投票数データと、推測される有権者人口をもとにした、各市区町村の投票率データを作成した。参議院議員選挙を用いた理由としては、1) 衆議院議員選挙と異なり、定期的に行われており、政局や世論の影響を受けづらいこと 2) 直近で行われた国政選挙であることという二点が挙げられる。なお、有権者人口については、後述する方法で、令和 2 年度、4 年度住民基本台帳、2020 年度社会・人口統計体系を用い推定した。

### 2.3 モデルの選定

統計的因果探索を行う機械学習アルゴリズムとして Zheng ら (2018) <sup>(2)</sup> による NOTEARS (Non-combinatorial

Optimization via Trace Exponential and Augmented lagRangian for Structure learning) モデルを採用した。本モデルは多くの因果探索手法が抱える、探索対象とする変数の増加に伴った組み合わせ論的な計算量の急増に対して、DAG のもつ非巡回という特性を滑らかな関数として定義することにより、DAG の探索を連続最適化問題として定式化し、計算を効率化したモデルである。本研究の、多数の変数での因果グラフの推定という目的を達成するために、当モデルのもつ計算の高効率性と因果グラフの表現という特性が有用であると判断し、本分析に採用した。

### 3. データセットの加工

教育用標準データセット SSDSE-A と、令和 2 年度、4 年度住民基本台帳、2020 年度社会・人口統計体系を元に作成した投票率データを用いた。投票率データは以下の手順で作成した。

- 1) 各市区町村の令和 2 年度、4 年度住民基本台帳の人口をもとに、令和 2 (2020) 年度から令和 4 (2022) 年度の人口増加率を求める。
- 2) 2020 年度社会・人口統計体系の 0-17 歳人口と SSDSE-A の総人口の差から 2020 年度の段階での 18 歳以上の有権者人口を求める。
- 3) 1 の人口増加率を 2020 年度の有権者人口に乘算し推定する。

また、データ作成後、各変数の分布を確認し、平均値からの差の絶対値が標準偏差の 3 倍を超えた値を外れ値としてデータから除き、最小値が 0.1 を切る項目については、最小値を 0、最大値を 1 とする正規化を行った。その結果抽出された 1639 市区町村からランダムに 700 市区町村を抽出し因果グラフ探索を行った。

なお、以上の分析は全て Python によっておこなった。

### 4. データ分析の結果

変数間の関係は以下のグラフのようになった。これは、各要素間の関係を矢印で記したもので、太いほど関係性が強いことを示している。また、矢印の向きは因果関係の方向を示している。

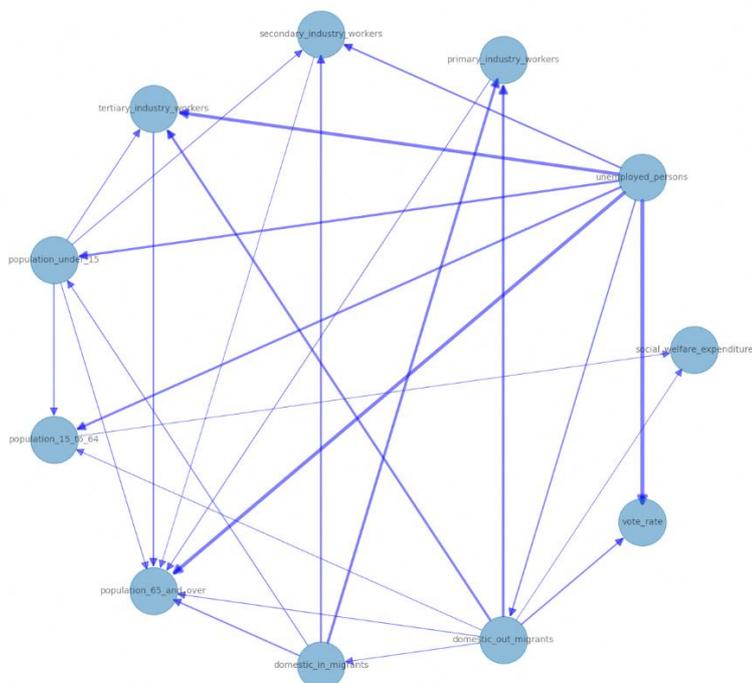


図 1 推定された因果グラフ

この図のように、投票率（vote\_rate）に関連するのは完全失業者率（unemployed\_persons）と、転出率（domestic\_out\_migrants）であることが分かった。

それでは、これらが実際に投票率にどのような影響を与えているかを見るために両者の関係を観察した結果が以下である。以下の表は、左軸の範囲に当たる市町村のなかで、投票率が該当範囲である市町村がどれほどの割合であったかを示している（表1）。

表1 完全失業者率、転出率の各区分における投票率区分の条件付き確率分布

失業率\投票率	上位25%	上位25～50%	下位25～50%	下位25%
上位25%	15%	19%	29%	36%
上位25～50%	12%	23%	29%	36%
下位25～50%	22%	30%	27%	21%
下位25%	51%	28%	14%	7%

転出率\投票率	上位25%	上位25～50%	下位25～50%	下位25%
上位25%	28%	29%	23%	21%
上位25～50%	21%	24%	25%	30%
下位25～50%	21%	23%	30%	27%
下位25%	30%	25%	23%	23%

以上の表を見ると、失業率が下がれば下がるほど投票率は上がることが分かる。また、転出率が極端に大きい・小さいほど投票率が上がることが分かった。

## 5. 結果の解釈

以上より、投票率と因果関係にあった変数は完全失業者率、転出率であった。一方で、転入率や年齢別人口分布、民生費の投資割合、各産業従事者率については、投票率との直接的関連が見られなかった。柿元ら（2004）<sup>(1)</sup>が指摘した民生費割合（行政投資額）、年齢別人口分布と投票率との相関は、転出率を交絡因子とした疑似相関であったと考えられる。同様に、投票率との関連が予想された各産業従事者率についても、転出率による疑似相関が生じていると推測される。

失業率の低下が投票率上昇をもたらすことは、柿元ら（2004）<sup>(1)</sup>の研究で示されるように、現在の景気の良さが過去の政治への肯定的評価、延いては未来の政治への期待につながるためだと考えられる。

また、転出率が投票率を変化させることは、転入率が投票率に影響しないことを考慮すると、投票率の変化要因が域内の人口量というより域内人口の質にあることを示していると考えられる。分析の結果より、転出率が極端に高い場合、低い場合それぞれについて、投票率を上昇させる構造が存在すると推測される。

まず、転出率が極端に高い地域で投票率が高くなるのは、激しい人口流出によって若年層が流出し、域内の高年層割合が高くなるためだと考えられる。一般的に高年層の投票率は高いため、若年層の転出により地域内

の高年者割合が増加することで地域全体の投票率が上昇すると推測される。

一方、転出率が極端に低い地域で投票率が高くなるのは、人口流動性の極端な低さによって住民間の緊密な人間関係が維持されるためだと考えられる。2022年度第26回参議院議員選挙における投票率上位の市区町村の特色は、例えば宮崎県西米良村<sup>(3)</sup>や山形県<sup>(4)</sup>で顕著なように、投票参加が家族や地域の中で形成された緊密な人間関係に織り込まれていることだ。こうした関係の中で選挙や投票は日常的话题の一つであり、不参加しづらい行事だという。しかし、緊密な人間関係という地域固有の文化は、域外への人口流出によって薄まってしまう。特に地域活動や文化の担い手である若年層の流出は痛手であり、文化を共有しない域外からの人口流入ですぐに回復するものではない。このように、文化を共有しているという質的特徴を持った住民の存在、そして彼らによる緊密な人間関係の維持が、投票率を上げるうえで重要だと考察される。

現行の施策の重点は、期日前投票所の設置や投票所への移動支援等の投票環境の整備と、主権者教育や選挙日程の周知等の啓発運動の二点に置かれている。本研究の分析の政策的含意は、投票率と地域文化の不可分な関係を指摘し、投票率向上施策の新たな方向性を示したことにある。例えば、家族や地域で選挙について話す機会を設けたり、祭り等の地域行事の振興を通して地域内コミュニティを支援したりする等の施策の実行余地があると考えられる。

本研究の分析の限界は二点ある。一点目は、本分析では選挙への関心や支持政党、地域への愛着などの心理学的変数を使用できなかったことだ。二点目は、データの制約から投票者人口の実測値を使用できず、推測値で代用したことだ。

## 参考文献

- (1) 柿元裕子、小柳昌子、西口香織里、今北恭平：“投票率の要因分析 —— 投票率上昇による民意反映を目指して”、ISFJ政策フォーラム2004研究報告（2004）。
- (2) Xun Zheng, Bryon Aragam, Pradeep Ravikumar, Eric P. Xing,：“DAGs with NO TEARS: Continuous Optimization for Structure Learning”、NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems、pp.9492 - 9503（2018）
- (3) 西日本新聞：“西米良村が投票率日本一 —— 決め手は「住民意識」の高さ”、2016年8月31日付西日本新聞朝刊。（2024年8月27日最終閲覧、<https://www.nishinippon.co.jp/item/n/271022/>）
- (4) 毎日新聞：“全国トップの高投票率 —— 山形県で、その理由を聞いてみた”、2022年6月16日掲載。（2024年8月27日最終閲覧、<https://mainichi.jp/articles/20220616/k00/00m/040/103000c>）