

Title	政治学における統計分析のフロンティア
Sub Title	Frontiers of data analysis in political science
Author	堀内, 勇作(Horiuchi, Yusaku)
Publisher	慶應義塾経済学会
Publication year	2015
Jtitle	三田学会雑誌 (Mita journal of economics). Vol.107, No.4 (2015. 1) ,p.545(1)- 552(8)
Abstract	現代政治学(ポリティカル・サイエンス)においては, 様々なかつ膨大な数量データと, 統計的手法を駆使した, 政治に関する諸仮説の検証が活発に行われている。本講演では, 「データ革命」とも呼ぶべき劇的な変化が起きている「サイエンス」としての現代政治学を, 具体例に基づいて紹介したい。また, 新しい時代に大学で何を教えるべきか, また学ぶべきかについても考察したい。
Notes	経済学講演会
Genre	Journal Article
URL	<a href="http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-20150101-0001">http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-20150101-0001</a>

経済学講演会

## Frontiers of Data Analysis in Political Science

堀内 勇作(Yusaku Horiuchi)

現代政治学（ポリティカル・サイエンス）においては、様々なかつ膨大な数量データと、統計的手法を駆使した、政治に関する諸仮説の検証が活発に行われている。本講演では、「データ革命」とも呼ぶべき劇的な変化が起きている「サイエンス」としての現代政治学を、具体例に基づいて紹介したい。また、新しい時代に大学で何を教えるべきか、また学ぶべきかについても考察したい。

### Abstract

Political scientists today investigate a range of hypotheses about politics and international relations using various types of statistical methods, as well as large amount of quantitative data. This seminar introduces such contemporary political science where a dramatic change, which should be called a “data revolution”, is occurring. In addition, it also discusses what should be taught and learned at universities in this new era.

## 政治学における統計分析のフロンティア\*

堀内 勇作

### 要 旨

現代政治学（ポリティカル・サイエンス）においては、様々なかつ膨大な数量データと、統計的手法を駆使した、政治に関する諸仮説の検証が活発に行われている。本講演では、「データ革命」とも呼ぶべき劇的な変化が起きている「サイエンス」としての現代政治学を、具体例に基づいて紹介したい。また、新しい時代に大学で何を教えるべきか、また学ぶべきかについても考察したい。

### キーワード

政治学、国際関係、統計学、ビッグデータ

### はじめに

本講演では、「データ革命」とも呼ぶべき劇的な変化が起きている「サイエンス」としての現代政治学を、具体例に基づいて紹介したい。また、新しい時代に大学で何を教えるべきか、また学ぶべきかについても考察したい。

### ポリティカル・サイエンスとは

私が政治学者であることを知人や友人に伝えると、しばしば、以下のような質問を受ける。まず、「次の選挙で自民党は勝つと思いますか」次に、「安倍政権は、どれくらいもつと思いますか」そして、最もよく聞かれる質問が、「政治家になりたいのですか」というものである。多くの人（特に日本人）にとって政治学者という職業は、現在進行中の日本政治の状況を論じたり、選挙結果を占

---

\* 私の母校訪問をアレンジして下さった藤原一平教授、講演のビデオ撮影、講演メモ作成などをして下さった藤原ゼミの皆様、そして多くの質問をして下さった参加者の皆様に感謝の意を表します。〔編集者註〕本稿は 2014 年 7 月 10 日に慶應義塾大学三田キャンパス東館 6 階 G-Sec Lab にて開催された経済学会・経済学部共催の講演内容をまとめたものである。

たり、政策過程に具体的に関与したりしている人というイメージがあるようである。しかし、「ポリティカル・サイエンス」と呼ばれる分野に取り組む学者の仕事は、そのようなイメージとは必ずしも同じではない。

ポリティカル・サイエンティスト（＝政治学者）が日々行っている作業は、様々な、かつ膨大な数量データと、統計的手法を使った、政治に関する諸仮説の検証である。データを用いると、政治に関して多くの人が当たり前だと思っていたことが、実は「そうではない」と分かることがある。つまり、「表と思われていたこと」（＝固定観念）を「裏にする」（＝覆す）ことがあるのである。一方、データを用いたからこそ、政治に関して多くの人があまり知らなかったこと（噂程度にしか聞いていなかったこと）が、実は「そうである」、つまり「裏」事情を「表」に出すこともある。政治学者が取り組んでいるのは、そのような「オセロ」のような地道な作業である。その取り組みを通じて、政治に関する真実を追求することが、サイエンスとしての現代政治学である。

では、どのような研究テーマが「オセロ」の題材となるのか。以下は、著者自身あるいは別の研究者が検証してきた「仮説」、すなわち「当たり前だと思われていること」あるいは「噂程度にしか聞いていなかったこと」、である。①低い投票率は、政治不信の表れである。②最高裁判所の判決は、政治的に中立である。③「ねじれ国会」においては、法案が通りにくい。④日米安保は、戦後経済成長にとっての重要な基礎である。⑤地方選挙の前に、投票を目的とした違法な住民票の移動が起きている〔文献1〕。⑥アメリカでは、民主党支持者と共和党支持者の二極分化が進んでいる。⑦途上国への援助は、被援助国の援助国に対する好感度を高めることに貢献している〔文献2〕。⑧貧困はテロリズムの温床である。⑨石油、天然ガス等の資源を豊富にもつ国（例：中東諸国）は、民主化できない。⑩民主主義国ほど、経済成長率は高い。ここでは一つひとつのテーマに関する研究結果を紹介することはできないが、このような仮説をデータで客観的に検証していくと、いろいろと見えてくることがある。それが現代政治学であり、その面白さである。

### 政治学では、どのような統計データが使われているか

では、これらの仮説を検証する上で、どのようなデータが使用されているのか。著者が博士号を取得した2001年前後の時期までは、政治学における統計分析というと、ほとんどの場合、世論調査やエリート調査などの質問票に対する回答データや、国・地域別、都道府県別、選挙区別あるいは市区町村別に集計されたデータ（世界銀行や国連が公表するデータ、国勢調査、地方財政統計、選挙の記録、等）が使われていた。

ところが、この10年間で政治学者が収集し分析するデータは、テクノロジーの急速な進歩にも伴い、劇的に変化した。このような「データ革命」とも呼ぶべき激変を経て、今日の政治学者は、無作為割当実験に基づくデータ、インターネット上のあらゆるデータ、デジタル化されたテキスト・

データをはじめとし、紛争・戦闘地域で入手したデータ、人工衛星で撮影した詳細データ、古今東西の地理情報データ、機能的磁気共鳴画像、自動顔認証技術や視線追跡技術を用いたデータ等、ありとあらゆるデータを駆使した研究を行っているのである。

データを解析するテクノロジーそのものも急速に進歩している。10年ほど前までは、クロス表分析、重回帰分析など、計量経済学の教科書（それも比較的初級から中級レベルの教科書）にある方法が使われていた。しかし今では、実験手法や統計モデルを精緻化することに（統計学者や経済学者ではなく）政治学者自身が貢献したり、コンピュータ・サイエンスの研究者とのコラボを通じてこれまで分析できないと思われてきた問題を追求したりもしている。また、数字では計測ができない質的（qualitative）データを、量的（quantitative）分析に体系的に組み合わせることにも、多くの学者が取り組んでいる。

その結果、以前は政治学（特に数理・計量的な研究）というと、応用経済学の一部であるかのように認識されることもあったが、今ではその認識が正しいとは言い難くなっている。現代政治学は、心理学、地理学、脳科学、歴史学、コンピュータ・サイエンスなど、様々な分野と融合しながら、政治に関する様々な仮説を厳密に検証することが求められているのである。

そのような「データ革命」の結果行われるようになってきた研究を、いくつか紹介したい。一つ目〔文献3〕は、著者の同僚らが行い、ワシントン・ポスト紙のブログで分析結果を紹介し話題になった研究である。ロシアのクリミア侵攻が世界的に話題になっている時、著者の同僚らは、アメリカの一般市民を対象としたインターネット世論調査を行った。この調査で世界地図を回答者に提示し、ウクライナの位置をコンピュータ上でクリックしてもらったところ、正解率はわずか6分の1であることが分かった。更に、回答者がクリックした地点と正解からの距離が遠いほど、アメリカの軍事介入を支持する確率が高いことも判明した。このように地図上の位置（緯度経度）データを、グーグル・マップ等を活用して分析する研究は、ここ数年急速に増えている。なお、この研究は方法論的に面白いだけでなく、これまでの常識を覆す可能性を示唆している。アメリカの軍事行動に関してアメリカ市民が支持するか否か、またその理由は何か。それらの問いに対して、世論調査を用いた研究が多く発表されてきた。その研究が示す一つの仮説は、アメリカ市民は、軍事行動の便益と負担を戦略的に考慮した上で、調査の質問に答えているというものである。しかし、著者の同僚らの研究は、実は世界情勢を良く理解していない人ほど、短絡的に軍事介入を支持する傾向が高いことを示唆している。

二つ目〔文献4〕の例は、この講演を準備している最中に国際関係論のトップ・ジャーナルに発表された論文である。政治体制、つまり民主主義であるか否かが、経済発展・経済成長にどのような影響を与えているかについては、政治学者と経済学者による膨大な研究の蓄積がある。他の影響をコントロールする限り、民主主義国ほど経済成長率が高いという研究もあれば、中国などの例を参照しながら、独裁国家の方がむしろ経済成長を促すことに成功しているという研究もある。その論

争に対して新たな貢献をすべく、同論文の著者らは、世界中の夜間における人口衛星写真を活用し、「明るさ」の変化率と国内総生産（GDP）の変化率の相関を調べた。その結果、多くの国において二つの変数の間に強い相関が確認されるものの、独裁国家では国内総生産が過大報告されている傾向があることが判明した。その乖離の程度がとりわけ激しいのが、中国とミャンマーである。このことは、日本の戦後高度経済成長に匹敵する成長を実現しているとされる中国の公式統計が、かなり疑わしいことをも意味している。非民主主義国ほど経済成長率が高いというのは、単に、データ上のエラーに過ぎない可能性があるのだ。

三つ目〔文献5〕の例は、ここ数年話題になってきている、遺伝学や脳科学の研究手法を政治学に応用したものである。アメリカの研究グループが、ギャンブルゲームに参加している82人の被験者の脳活動を計測した結果、民主党支持者と共和党支持者の脳活動が異なることが明らかになった。これは脳活動のデータに基づく政党支持態度を予測する研究であるが、同研究グループの主張によると、別の研究グループによって進められている遺伝情報に基づく政党支持態度の予測よりも予測精度が高いらしい。このような脳活動データや遺伝子データを用いる政治学研究に関しては、ここ数年、反論、再反論が活発に出されている。著者自身は、これらの研究の今後の発展の可能性は必ずしも高くないと考えているが、ここでは、社会科学だけでなく自然科学の研究にも取り組もうとする現代政治学の例として紹介したい。

四つ目〔文献6〕の例は、コンピュータ・サイエンスの領域で発展している機械学習（Machine Learning）の応用例である。ハーバード大学のグループは、あるブログ上に記入された、アメリカ大統領候補者とされる政治家らに対する膨大な数のコメントを分析した。具体的には、当時の民主党のリーダー（現国務長官）であるジョン・ケリーが2006年11月に失言したことで、同候補者に対する支持、不支持の割合がどう変化したかを計測することを試みた。では、どのようにして、「定性」的なコメント内容を「定量」化したのか。同グループは、まず膨大な数のコメントから無作為に抽出された数百のコメントを、複数のリサーチ・アシスタントに丁寧に読んでもらい、内容に基づいてコメントをグループ分けしてもらった。その後、コンピュータに人間によるコーディングの方法を学習させ、解析されたコーディング・メカニズムによって残りのコメントを自動的にコーディングさせたのである。この結果、ジョン・ケリーに対する不支持は失言によって激増したことが統計的に示された。なお、ケリーはこの失言が一つのきっかけとなり、2008年の大統領民主党予備選挙では初期の段階でレースから脱落したとも言われている。

このような膨大なテキスト情報を体系的に分析する政治研究は、ここ数年目覚ましく進歩している。これらの研究は、従来の政治学が使っていたメソッドの範疇を越えて、コンピュータ・サイエンスや、「デジタル・ヒューマニティ」と呼ばれる領域とも融合しながら、様々な仮説の検証を行っている。分析で使われるテキスト情報も、非常に多岐にわたっている。毎日毎時記録されるブログやツイッターなどのコメント、様々なウェブサイト上にある情報、更にはデジタル化された書籍、論

文、資料、政府刊行物など、ありとあらゆる文字情報は分析対象となりうる。ウェブサイト上にある天文学的な量の文字情報を、コンピュータを用いて自動的にダウンロードし、分析可能で体系的なデータセットに変換する技術（ウェブ・スクレイピング）も、政治研究に活用されるようになってきている。例えば、オーストラリアの連邦議会図書館では、選挙運動期間中の政党、候補者のウェブサイトをもろごとダウンロードして保存している。オーストラリア連邦議会における全発言記録とあわせて、政治家の発言内容がどのような要因によって規定されているかを分析する研究を、著者自身、目下準備中である。

コンピュータによるテキスト分析は、いわゆる「ビッグデータ」分析の典型であるが、その例を一つ紹介したい。これは、政治学における計量分析方法論の大家であるハーバード大学のゲイリー・キング教授のチームが行い、ここ1～2年ほど世界中で話題になっている研究〔文献7〕である。日本語でも、日経ビジネス・オンライン（2014年2月3日）に紹介されているので参照されたい。この研究でキング教授のチームは、中国のソーシャル・メディアに記録された投稿を、中国政府が検閲する前に超高速でダウンロードした。その後、クラウド・コンピュータ上に記録された膨大な量のビッグデータをコンピュータに内容解析させ、どのような投稿が中国政府によって削除される傾向が高いかを調べた。その結果、中国政府が最も恐れているのは、団体行動を呼びかける投稿であることが判明した。中国政府は、政府批判そのものよりも、団体行動に過敏に反応していることがデータによって示されたのである。

以上、地理情報、人工衛星画像、脳活動データ、膨大なテキスト情報を活用した最新の政治学研究をいくつか紹介してきた。次に、著者自身が行った研究を一つ紹介したい。

### 有権者が投票先を決める際に何を判断材料にしているか

この研究〔文献8〕は、選挙における投票行動に関するものである。従来、政治学の教科書では、有権者が投票先（候補者、政党）を決める際、政党の政策内容、候補者の資質、政権党の政策実績、等に基づいて投票すると書かれてきた。また、有権者はそのような情報に基づいて投票する「べき」であるし、実際にそのような情報に基づいて投票している「はず」である、と考えられてきた。ところが、多くの研究者が新しいデータを用いた分析をしてみると、実は、有権者はトリビアな（取るに足りない）情報に基づいて投票している傾向もあることが分かってきた。「つまらない」要因に基づいて有権者が投票しているとすれば、それを調べることは「つまらない」研究ではない。民主主義のあり方、国の政策過程の問題を考える上で、極めて重要な意味がある。

そこで著者らのチームは、自動顔認証技術を用いて各候補者の笑顔度を測定し、客観的に計算された笑顔度の選挙結果に与える影響を推計した。具体的には、2000年衆議院議員総選挙（日本）に立候補した候補者ポスターにある顔写真と、2004年のオーストラリア下院議員選挙に立候補した候



補者が使用した「How to Vote Card」にある顔写真を用いて笑顔度を測定し、他の様々な要因をコントロールした上で、各選挙における候補者の笑顔度が得票率に与えた影響を推計した。

推計結果を紹介する前に、この研究で用いた自動顔認証技術について説明したい。自動顔認証は近年技術革新を遂げて、様々な商品に応用されてきている。例えば、スマートフォンで写真を撮ろうとした時に、被写体に人物の顔があるかを瞬時に判別してくれるのは、自動顔認証技術が組み込まれているからである。デジタルカメラにも、アップル社やグーグル社などの写真整理・編集ソフトにも、この技術が用いられている。

著者のチームが使用した技術は、オムロン社が開発した Okao Vision というものである。この研究を実施したのは既に5年ほど前で、当時は笑顔だけを数値化できたが、オムロン社の技術開発は更に進み、現在では人間の五つの感情を示す表情を数値化することに成功しているとのことである。ちなみに、オムロン社が技術開発を行う上で、最初に笑顔の数値化に取り組んだ理由が興味深い。古今東西、世界には様々な人種、民族、文化的背景をもつ人々がいるが、「ハッピー」な時に示される顔の特徴は万国共通らしい。それを踏まえて、比較的数値化しやすい笑顔を測定することに挑んだとのことである。

さて、その Okao Vision を用いて各候補者の笑顔度を測定してみると、100%の笑顔は、オーストラリア前首相のジュリア・ギラードさんと現参議院議員の松沢成文さんが該当した。笑顔度が50%だと、笑っているのか笑っていないのか分からない微妙な顔になる。日本では、現衆議院議員の野田聖子さんが該当した。ちなみに笑顔度が0%の政治家の一人は、現衆議院議員の小沢一郎さんであった。

そもそも日本とオーストラリアでは立候補者の笑顔度に違いがあるのか、平均値を計算してみたところ、日本は50%、オーストラリアでは73%であった。オーストラリアの政治家は、日本に比べかなり笑顔度が高いのはなぜか。これは著者らの統計分析からだけでは分からないが、文化的違いもあると思われる。例えば日本では、女性は大笑いせず、笑う時は口を手で覆い隠すことが美德とされている。一方、「No worries, mate」とよく口にするオーストラリア人の間では、歯を見せることなど気にせず大きく笑うことが相手に好印象を与えるとされているのかもしれない。

各候補者の笑顔度と得票率の関係を統計的に推計した結果、日本では、笑顔度が0%から100%になると得票率が2.3%ポイント上昇することが分かった。この影響はオーストラリアでは更に顕著で、5.2%ポイントも上昇する。どちらの影響も統計的に有意であり、影響の大きさも無視できない。特に、オーストラリアでは接戦の小選挙区が多いので、他の影響を十分にコントロールした上でも5%ポイントもの影響があるということは、候補者の笑顔によって勝敗が左右されかねないことを意味している。

では、なぜ笑顔度が高い候補者は高い得票率を得るのか。著者らが様々な分野における先行研究をレビューする中で、脳心理学における興味深い論文を知った。機能的磁気共鳴画像 (functional



magnetic resonance imaging) を用いたある研究によると、笑顔を見せられた被験者は、綺麗な音楽を聴いたり、美味しいものを食べたり、快適な空間にいたりする時に活動する脳の部位が、同じように活動するらしい。また、笑っている人物の名前の方が、笑っていない人物の名前よりも覚えやすいという実験結果も別の研究で示されている。これらの研究によれば、笑っている候補者を見ると、その人物が有能であるか否かを合理的に判断することもなく、無意識のうちに脳の部位が反応し、名前を覚えてしまうのである。選挙において候補者は、名前を覚えてもらうために顔写真を使うので、これらの研究が示すことは著者らが得た計量分析結果と整合的である。重要な点は、候補者に政治家としての能力があるか否かとは関係なく、有権者が投票先を判断してしまっている可能性があるということである。政治家の側からすれば、笑った方が選挙で勝つためには良いということになるが、有権者にとっては、笑っている政治家が有権者を幸せにするとは限らないのである。

### 新しい時代に何を教えるべきか、何を学ぶべきか

冒頭で述べた通り、この10年間で、社会科学において用いられるデータや分析手法は大きく変化してきている。そうした中、何を大学で教えるべきか、何を学生は学ぶべきか。著者は、新しい時代に対応できる次世代の学生を育てていく必要性を痛感し、目下、ダートマス大学の同僚たちと、「数量社会科学」に関する教育プログラムの立ち上げを準備している。

次世代の学生に必要とされるものは何か。まずは、自由な発想であろう。発想が豊かでないと、どんなにビッグデータがあっても何も生まれてこない。そして、自由な発想をカタチにできる技術も必要だ。その両方が、これからの学生には必要不可欠であると、著者は考えている。

では、自由な発想はどこから生まれてくるのか。一見すると自分とは「関係ない」分野と思える他分野のことも積極的に学ぶ姿勢が、新しい発想を生むと思っている。自分の専門外の人の話しを聞くことも重要であろう。そして、世界中のどこにでも行ってみようというエネルギーも必要だ。色々な人、知らないものに出会った時こそ、新しく、面白い発想が生まれると、著者は常に思ってきた。

発想をカタチにする技術を得るには何が必要か。統計学は必須である。しかし、膨大なデータを集めて分析をする上では、統計学だけでは足りなくなっているのも確かである。コンピュータ・サイエンスの基礎知識と言語を学ぶことも、社会科学を学ぶ学生に必要なようになってきている。プレゼンテーション・スキルも重要である。膨大なインターネット上のデータをネットワーク図で示したり、世界の様々な状況を地図で示したり、文章情報をワード・クラウドで示したりすることは、これから益々増えていくであろう。つまり、データを、数字だけではなく、誰にでも分かりやすくビジュアルに示すことが要求されてきていると思われる。そして最後は外国語。外国語ができると、たくさんのところに行き、たくさんの人の話しを聞くことができる。ダートマス大学では、大学に入って習い始めた言語を、卒業までに使えるようにしている学生に会うことは、全く珍しくない。4

年間に二つか三つの言語を習得し、世界各地でインターンシップやボランティアに参加する学生も多くいる。日本においても、これからの学生は、英語は勿論、中国語やフランス語などを習得していく必要があるだろう。

(ダートマス大学政治学部三井冠准教授)

#### 参 考 文 献

- [1] Horiuchi, Yusaku, and Kentaro Fukumoto. 2011. "Making Outsiders' Votes Count: Detecting Electoral Fraud through a Natural Experiment." *American Political Science Review* 105 (3): 586–603.
- [2] Goldsmith, Benjamin E., Yusaku Horiuchi, and Terence Wood. 2014. "Doing Well by Doing Good: The Impact of Foreign Aid on Foreign Public Opinion." *Quarterly Journal of Political Science* 9 (1): 87–114.
- [3] Dropp, Kyle, Joshua D. Kertzer, and Thomas Zeitzoff. 2014. "The Less Americans Know about Ukraine's Location, the More They Want U.S. to Intervene." *Monkey Cage, The Washington Post*, April 7. URL: <http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/04/07/the-less-americans-know-about-ukraines-location-the-more-they-want-u-s-to-intervene/>
- [4] Magee, Christopher S. P., and John A. Doces. 2014. "Reconsidering Regime Type and Growth: Lies, Dictatorships, and Statistics." *International Studies Quarterly*, forthcoming.
- [5] Darren, Schreiber, Greg Fonzo, Alan N. Simmons, Christopher T. Dawes, Taru Flagan, James H. Fowler, and Martin P. Paulus. 2013. "Red Brain, Blue Brain: Evaluative Processes Differ in Democrats and Republicans." *PLoS ONE* 8(2): e52970. doi:10.1371/journal.pone.0052970
- [6] Hopkins, Daniel J., and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* 54 (1): 229–247.
- [7] King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107 (02): 326–343.
- [8] Horiuchi, Yusaku, Tadashi Komatsu, and Fumio Nakaya. 2012. "Should Candidates Smile to Win Elections? An Application of Automated Face Recognition Technology." *Political Psychology* 33 (6): 925–933.