

# 世論調査と出口調査データから推定した 傾向スコアを用いた選挙予測について<sup>†</sup>

福田 昌 史\*

Predicting election results using propensity scores estimated based on  
pre-election survey data and exit poll data

Masafumi Fukuda\*

報道機関では、選挙に関する調査を主に2種類行っている。1つは投票日の1週間前に行う情勢調査で、もう1つは投票日当日に行う出口調査である。これらの調査での「誰に投票するか(したか)」の回答を見ると、情勢調査に比べて出口調査の結果の方が実際の開票結果に近い。本稿では、この2種類の調査の回答データから傾向スコアを推定する。推定には、性別、年齢、投票区の市郡別、政党支持、投票の基準の回答データを共変量として使った。そして推定した傾向スコアを使って3つの方法(A)傾向スコアで層別して加重平均、(B)(A)の係数を修正したもの、(C)傾向スコアを使った重み付け、によって情勢調査からの予測の改善を試みた。その結果、(B)を使った方法において、有力候補の7割の予測値に改善が見られた。

A newspaper company usually conducts two kinds of surveys concerning an election. One is a pre-election survey done a week before the election for the purpose of predicting election results. Another is an exit poll done just after the election for the purpose of obtaining information about the relationship between voters' demographic characteristics, their political attitudes, and their electoral decisions. It is known that percentages of answers to the question, "Which candidate will you vote for?" are further away from the actual election outcome than those obtained through exit polls. In this paper, we estimated propensity scores through both kinds of survey data. The estimation was performed by using covariates such as sex, age, residential area of the voter, political party identification and criterion for selecting the candidate. Next, three adjustment methods were applied to improve predictions using pre-election survey data: (A) weighted averaging by splitting estimated propensity scores into strata, (B) a modified method of (A), and (C) propensity weighting. As a result, forecasted shares of the vote were improved for 70 percent of the candidates in the sample using method (B).

*Key Words and Phrases:* 選挙予測, 情勢調査, 出口調査, 傾向スコア

## 1. はじめに

報道機関では選挙に関連する調査として、主に選挙情勢世論調査(以降、単に情勢調査と呼ぶ)と出口調査を実施している。情勢調査は、その選挙の情勢を把握し、報道するために投票

\* 中央大学理工学研究科, 〒112-8551 東京都文京区春日 1-13-27

毎日新聞東京本社世論調査室, 〒100-8051 東京都千代田区一ツ橋 1-1-1

<sup>†</sup> 本論文の内容は、執筆者個人の実験的な分析であり、毎日新聞社が実際に使っている手法ではありません。

日一週間前の週末に行われる。調査では「誰に（どの政党に）投票するか」を質問し、候補者や政党の名前が挙がった割合をもとに予想得票率を算出して、各候補者の当落予測や政党の議席数予測を行っている。この調査結果には、標本誤差だけでなく、投票する候補者をまだ決めていない回答者がいることによる誤差や、投票しようとしている候補者を答えているにもかかわらず投票に行かない回答者が含まれることによる誤差など、さまざまな要因からのノイズが含まれており、選挙予測ではこのようなノイズや情勢調査特有の偏りをいかに補正するかが重要である。

出口調査は、当落判定や投票行動の分析記事のための参考データ収集を目的に行われる。この調査は投票終了時刻の直前まで実施しており、時間的な順序関係から事前の予測には直接使うことができない。しかし、情勢調査の結果より開票結果に近い数字となっていることから、情勢調査と出口調査で共通な調査項目を媒介にして情勢調査の結果を出口調査の結果に近づけることができれば、情勢調査結果を修正して開票結果を予測する新しい方法が考えられる。

Taylor (2001) は、Rosenbaum and Rubin (1983) が提案した傾向スコアを応用し、インターネット上でパネルを対象に調査した結果を電話調査の結果に近づける手法を使うことによって2000年の米国大統領選挙結果を予測し、傾向スコアの調査データ補正への応用可能性を示した。

本論文では、傾向スコアを用いて情勢調査の結果を出口調査の結果に近づけることによる選挙予測の改良法を提案する。以下第2節では、情勢調査と出口調査の概要と、従来の選挙予測について説明する。第3節ではそれらの調査データから推定される傾向スコアを使った補正の手法について述べ、第4節では実際のデータにあてはめた分析例とその結果を示し、第5節でまとめを述べる。

## 2. 選挙調査の概要と予測

### 2.1 情勢調査と出口調査の概要

国政選挙や注目度の高い知事選、政令市長選では、報道機関によって情勢調査や出口調査が行われている。以下ではそれらの概要について説明する。

情勢調査は、主に投票日1週間前の週末に2日から3日かけて電話調査により行なわれる。かつては選挙人名簿から調査対象者を抽出し、対象者の電話番号を調べ、番号が判明した対象者に電話調査を行っていたが、近年ではRDD (Random Digit Dialing) 方式による調査が中心となっている(報道機関の情勢調査については、松田, 2002, 窪田, 2002, 長江, 2002)。標本サイズは、選挙区の情勢や同時期に実施する調査の規模により変動する。質問項目としては性別、年齢などの基本属性の他に、投票に行くかどうか、誰に投票するつもりか、支持する政党や選挙の争点に関する質問がある。そして「誰に投票するか」の回答を集計し、適当な補正を経て得票率を予測する。

情勢調査の回答者の中で、投票する候補を、「まだ決めていない」という回答者は、2割から4割いる。そして、「決めていない」回答者を除いた、投票する候補の名前を挙げた回答者の割合を、挙名率と呼ぶ。これは、調査結果の信頼性を見るための指標の1つであるが、選挙予測の当たり外れとの相関はそれほど高くなく、当選を的中させた選挙区の平均挙名率と外した選挙区の平均挙名率は、ともに0.7程度である。

出口調査は、選挙区を市部・町村部に分けた層別抽出によって30から60の投票所を抽出し、ひとつの投票所でおおよそ50人から80人分の回答を集めるという層別二段抽出を採用している。調査対象者は、投票を終えて投票所から出てくる人の中から順番に一定の人数間隔で抽出される。その人数間隔は投票所ごとに過去の選挙から算出した推定投票率と、その投票区の選

挙人名簿登録者数から決められる。調査で使用する質問票には、性別、年齢、誰に投票したか、支持政党などの質問が印刷されており、これを手渡し、対象者自身に回答を記入してもらう。出口調査においても、誰に投票したかを尋ねる質問への回答を得られないケースが平均で5%程度あるが、ほとんどの場合、調査結果は開票結果と近い数字となる。

## 2.2 調査結果からの選挙予測

予測の対象とする選挙の有権者数を  $N$  人とする。そして、その選挙で予測の対象とする候補者を A と呼ぶことにする。 $v_i$  を  $i$  番目の有権者がその選挙で投票に行くかどうかを表す変数

$$v_i = \begin{cases} 1 & \text{投票に行く場合} \\ 0 & \text{それ以外} \end{cases} \quad (i=1, \dots, N)$$

とし、 $Y_i$  を投票に行った  $i$  番目の有権者が候補者 A に実際に投票するかどうかを表す変数

$$Y_i = \begin{cases} 1 & \text{候補者 A に投票する場合} \\ 0 & \text{それ以外} \end{cases} \quad (i=1, \dots, N)$$

とする。簡単のため、無効票や案分票を考えないことにすると、候補者 A の得票率は、

$$\bar{Y} = \frac{1}{\sum_{i=1}^N v_i} \sum_{i=1}^N v_i Y_i$$

と書ける。

その選挙で情勢調査が行われた場合、 $z_i$  を、 $i$  番目の有権者が情勢調査の対象者に選ばれ、調査に協力し、誰に投票するか回答したかどうかを表す変数

$$z_{1i} = \begin{cases} 1 & \text{情勢調査で誰に投票するか回答した場合} \\ 0 & \text{それ以外} \end{cases} \quad (i=1, \dots, N)$$

とする。そして、 $y_{1i}$  を、情勢調査に回答した  $i$  番目の有権者が候補者 A に投票するつもりかどうかを表す変数

$$y_{1i} = \begin{cases} 1 & \text{候補者 A に投票すると回答した場合} \\ 0 & \text{それ以外} \end{cases} \quad (i=1, \dots, N)$$

とすると、候補者 A の予測される得票率は、

$$\bar{y}_1 = \frac{1}{\sum_{i=1}^N z_{1i}} \sum_{i=1}^N z_{1i} y_{1i}$$

である。出口調査の場合も同様に記号を定義し、 $z_{0i}$  を、 $i$  番目の有権者が選挙に行き、出口調査の対象者に選ばれて、投票した候補者を回答したかどうかを表す変数

$$z_{0i} = \begin{cases} 1 & \text{出口調査で誰に投票したか回答した場合} \\ 0 & \text{それ以外} \end{cases} \quad (i=1, \dots, N)$$

とし、 $y_{0i}$  を、出口調査に回答した  $i$  番目の有権者が候補者 A に投票したかどうかの回答を表す変数

$$y_{0i} = \begin{cases} 1 & \text{候補者 A に投票したと回答した場合} \\ 0 & \text{それ以外} \end{cases} \quad (i=1, \dots, N)$$

とすると、出口調査から予測する候補者 A の得票率は、

$$\hat{y}_0 = \frac{1}{\sum_{i=1}^N z_{0i}} \sum_{i=1}^N z_{0i} y_{0i}$$

である。

通常、情勢調査の結果  $\hat{y}_1$  と実際得票率  $\bar{Y}$  の間には強い相関関係がある。図1は、過去の衆議院議員選挙における、ある政党の候補者の情勢調査結果と実際の開票結果をプロットしたものである。それぞれの点は、その政党の公認候補の情勢調査結果（水平軸）と開票結果の得票率（垂直軸）を表している。図では、単相関係数は0.94である。そして、 $\bar{Y}$  と  $\hat{y}_1$  の間に回帰モデル  $\bar{Y} = g(\hat{y}_1)$  を仮定して得票率の予測値の補正を行う（西平, 1984 など）。回帰モデル  $g$  として  $\hat{y}_1$  の3次式を使っている例もあるが、単回帰で十分である場合が多い。

国政選挙における予測の場合、回帰モデルは政党ごとに異なり、接戦が予想される選挙区では、情勢調査の結果をそれぞれの政党の回帰モデルに当てはめることによって順位の逆転があり得るかどうかを予想することができる。しかし、知事選などの地方選挙では候補者の政党色が薄く、また2003年の統一地方選挙では無党派層へアピールするために政党の推薦を拒否する有力候補が現れて話題となったことなどを考えると、政党別の回帰式を作ることができず、全ての候補者について単一の回帰式を当てはめることになり、勝敗を予測するにはあまり有効な方法とはならず、何らかの別の予測手法を考える必要が出てくる。

### 3. 調査データの傾向スコア補正による選挙予測

#### 3.1 傾向スコアと補正のねらい

$\bar{Y}$  と  $\hat{y}_1$  の関係を  $g$  というモデルで表し、新しく得られた情勢調査のデータを  $g$  に当てはめて選挙予測するのが従来のやり方だが、出口調査の結果  $\hat{y}_0$  が  $\bar{Y}$  と近い値である場合、 $\hat{y}_0$  を  $\hat{y}_1$  から予測する方法  $g'$  を確立すれば、結果的に情勢調査から開票結果を予測できる可能性がある。

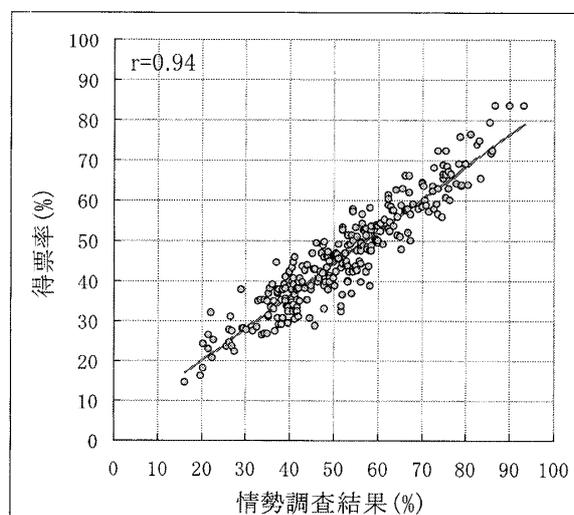


図1 情勢調査結果（国政選挙の場合）

出口調査のデータからは、公式に発表される開票結果からは知ることのできない「どのような属性・意識の人が誰に投票したか」という情報を得ることができる。以下では、これらの付加的な情報を傾向スコアを通じて予測に活かす方法を検討する。なお、ここでは仮に同じ回答者が情勢調査と出口調査の両方に回答したとき、調査の方式が異なっても同じ質問に対しては同じ回答をすると仮定しておく。

傾向スコアによって、インターネット調査の結果を電話調査などの従来型調査結果に近づける Taylor (2000) や星野 (2003, 2004) の試みは、性別、年齢などの人口統計学的属性だけでなく意見や行動に関する測定値を傾向スコアという1次元の数値に置き換え、この数値をサンプルへの重み付けに利用して抽出の偏りを調整するものである。ここでの傾向スコアの値は、共変量に使われる質問への回答パターンがどちらの調査に現れやすいかという数値(確率)となる。

いま、分析の対象とする集団を、情勢調査か出口調査のどちらかに回答している集団、つまり、

$$z_{1i}=1 \quad \text{または} \quad z_{0i}=1 \quad (i=1, \dots, N)$$

となる集団について考える。これは、どちらの調査でも観測することのできない  $z_{1i}=0$  かつ  $z_{0i}=0$  となるデータを考えないということであるが、出口調査の対象者が投票者集団を十分に代表しているのであれば問題にはならない。

また、実際に得られるデータについて、選挙区の有権者数  $N$  が大きい場合は、情勢調査と出口調査の両方の対象者になることはほとんどないと考えられ、つまり  $z_{1i}=1$  かつ  $z_{0i}=1$  となる  $i$  が出現する可能性はきわめて低いので、有権者  $i(i=1, \dots, N)$  は、

- $z_{1i}=1$  かつ  $z_{0i}=0$
- $z_{1i}=0$  かつ  $z_{0i}=1$

の2通りのうちのどちらかであるとして考える。そこで、2つの調査協力に関する変数  $z_{1i}, z_{0i}$  を、

$$z_i = z_{1i} = 1 - z_{0i}$$

という1つの変数に置き換える。つまり、

$$z_i = \begin{cases} 1 & \text{情勢調査の対象者} \\ 0 & \text{出口調査の対象者} \end{cases}$$

ということになる。

いま、個々の対象者の属性や意見などからなる共変量ベクトルを  $\mathbf{x}$  とする。どちらかの調査に回答する集団のなかで、共変量ベクトル  $\mathbf{x}$  を与えたときに情勢調査に回答している確率は、

$$e(\mathbf{x}) = \Pr(z=1|\mathbf{x})$$

で、これは傾向スコア (Rosenbaum and Rubin, 1983) と呼ばれる。この値はどのような属性・意識の人が情勢調査に答えやすいかに関係している数値である。そして、 $e(\mathbf{x})$  が0に近い人ほど投票に行き、その  $\mathbf{x}$  は出口調査の回答として出現しやすく、 $e(\mathbf{x})$  が1に近い人ほど情勢調査のみに回答し、その  $\mathbf{x}$  は出口調査の回答として出現しにくい、つまりあまり投票に行かない傾向があり、これを使って調査への参加の偏りを調整できると考えられる。

### 3.2 傾向スコアによる補正のための条件

Rosenbaum and Rubin (1983) では、傾向スコアが Balancing Score の 1 つであることが示されている。ここで Balancing Score とは、

$$\mathbf{x} \perp\!\!\!\perp z | b(\mathbf{x}) \quad (1)$$

が成立するような  $\mathbf{x}$  の関数  $b(\mathbf{x})$  である。つまり、それが与えられたとき、共変量  $\mathbf{x}$  と調査識別変数  $z$  が独立になる関数のことである。式(1)で  $b(\mathbf{x})$  を  $e(\mathbf{x})$  に置き換えると  $\mathbf{x} \perp\!\!\!\perp z | e(\mathbf{x})$  となる。これは、

$$\Pr(\mathbf{x}, z | e(\mathbf{x})) = \Pr(\mathbf{x} | e(\mathbf{x})) \quad (2)$$

ということであり、 $e(\mathbf{x})$  が与えられたときの  $\mathbf{x}$  の分布は  $z$  に依存しないということである。一般に傾向スコア  $e(\mathbf{x})$  は未知なので、推定する必要があるが、推定された傾向スコアによって式(2)のように共変量のバランスがとれているかどうかを確認する必要がある。

また、Rosenbaum and Rubin (1983) 定理3では、強く無視できる割り当て条件

$$(y_1, y_0) \perp\!\!\!\perp z | \mathbf{x} \quad (3)$$

を仮定すると、

$$(y_1, y_0) \perp\!\!\!\perp z | e(\mathbf{x}) \quad (4)$$

も成立することが示されている。つまり、 $\mathbf{x}$  が与えられたとき  $y_1, y_0$  の分布が  $z$  に依らないならば、同じ値の傾向スコアを持つ人であれば、投票に関する回答の分布が  $z=1$  (情勢調査回答者) の群と  $z=0$  (出口調査回答者) の群で同じということを意味する。ここでは、この強く無視できる割り当て条件を仮定して議論を進める。

### 3.3 補正方法の提案

得られているデータでは、情勢調査の対象者  $i$  の出口調査の回答  $y_{0i}$  や、出口調査の対象者  $j$  の情勢調査の回答  $y_{1j}$  は潜在的には存在するが未観測である。

いま、 $t_i$  を対象者  $i$  が仮に両方の調査に答えたときの回答の変化

$$t_i = y_{1i} - y_{0i} = \begin{cases} 1 & \text{(情勢調査では投票意志があったが、} \\ & \text{出口調査では他の候補に投票したと回答)} \\ 0 & \text{(情勢調査での投票意志の回答} \\ & \text{と出口調査での投票結果が一致)} \\ -1 & \text{(情勢調査では投票意志はなかったが、} \\ & \text{出口調査では投票したと回答)} \end{cases}$$

とする。

多くの有権者は、1週間前に投票する候補を決めていれば、そのまま意中の候補に投票すると考えられ、その場合は  $t_i=0$  であると仮定することができる。すると、 $E[t | \mathbf{x}] = E[y_1 - y_0 | \mathbf{x}] = 0$  であり、

$$E[y_0 | e(\mathbf{x})] = E[y_1 | e(\mathbf{x})] \quad (5)$$

も成立する。これは傾向スコアが等しいときは候補者 A への投票に関する回答の期待値は両方の調査において同じであるということである。そして、傾向スコアの推定値  $\hat{e}(\mathbf{x})$  を使用し、以下の3つの方法で得票率予測値の補正方法を試みる。

**方法 A：傾向スコアによる層別支持率の加重平均**

$t_i=0$  を仮定すると、式(5)と、式(4)から

$$\begin{aligned} E[y_0|e(\mathbf{x})] &= E[y_0|z=0, e(\mathbf{x})] = E[y_0|z=1, e(\mathbf{x})] \\ E[y_1|e(\mathbf{x})] &= E[y_1|z=0, e(\mathbf{x})] = E[y_1|z=1, e(\mathbf{x})] \end{aligned} \quad (6)$$

であるので、

$$E[y_0|z=0, e(\mathbf{x})] = E[y_1|z=1, e(\mathbf{x})] \quad (7)$$

である。これで情勢調査の回収サンプルから 1 週間後の出口調査の結果を推定することができ、

$$\begin{aligned} E[y_0|z=0] &= E_{e(\mathbf{x})}[E[y_0|z=0, e(\mathbf{x})]] \\ &= E_{e(\mathbf{x})}[E[y_1|z=1, e(\mathbf{x})]] \\ &= \int_0^1 E[y_1|z=1, e(\mathbf{x})] f(e(\mathbf{x})) d(e(\mathbf{x})) \\ &\approx \lim_{k \rightarrow \infty} \sum_{h=1}^k E[y_1|z=1, e(\mathbf{x}) = h/k] \hat{f}_0(h/k) \cdot \frac{1}{k} \\ &\approx \sum_{h=1}^{10} \bar{y}_1^{(h)} \cdot \frac{n_1^{(h)}}{n_1} \\ &= \sum_{h=1}^{10} w^{(h)} \cdot \bar{y}_1^{(h)} \equiv \bar{y}_1^{(A)} \end{aligned} \quad (8)$$

となる。ここで  $f_0(\cdot)$  は出口調査の回答者層の傾向スコアの密度関数、 $h$  は回答者集団を傾向スコアで長さ  $1/k$  の  $k$  個の区間で層別したときの層の番号、 $\bar{y}_1^{(h)}$  は  $h$  番目の層での候補者 A へ投票する回答者の割合、 $n_1$  は出口調査の対象者数、 $n_1^{(h)}$  は出口調査の対象者集団の中の  $h$  層の大きさである。

**方法 B：A の係数を修正し、加重平均を計算する**

方法 A だけでは予測しきれない部分を、重み付けを変えることによって補う。

$$y_1^{(B)} = \beta_0 + \beta_1 \cdot \bar{y}_1^{(1)} + \beta_2 \cdot \bar{y}_1^{(2)} + \cdots + \beta_{10} \cdot \bar{y}_1^{(10)}. \quad (9)$$

これは、 $t_i=0$  の条件を必要としない。この式の係数  $\beta_0, \beta_1, \dots, \beta_{10}$  は、重回帰モデルの係数として推定する。

**方法 C：傾向スコアを使った重み付け (propensity weighting)**

$z=0$  群のみの周辺平均の推定法、

$$y_1^{(C)} = \hat{E}[y_1|z=0] = \frac{\sum z_i y_i (1 - \hat{e}_i(\mathbf{x}))}{\sum \hat{e}_i(\mathbf{x})} \bigg/ \frac{\sum z_i (1 - \hat{e}_i(\mathbf{x}))}{\sum \hat{e}_i(\mathbf{x})} \quad (10)$$

により、出口調査の対象者が情勢調査に回答した場合の期待値を推定することで補正を行う。

**4. 調査データへの適用****4.1 使用したデータ**

ここで使う分析には、情勢調査と出口調査の両方で共通する質問の回答データを必要とする。質問の中には、支持政党のように常に共通して聞いているものもあるが、それぞれの選挙での争点に関する質問もあり、これは 1 回しか使われない。特に出口調査では全体の設問数が少ないので、両調査で共通して聞いている質問は限られてくる。そこで常に質問している項目である、「性別 (男性, 女性)」、「年齢 (20 歳代から 70 歳以上まで 6 カテゴリー)」、「支持政党」と、

表1 情勢調査と出口調査の挙名回収数

調査	挙名回収数	調査	挙名回収数
世論 1	469	出口 1	2080
世論 2	524	出口 2	3868
世論 3	641	出口 3	3828
世論 4	492	出口 4	2852
世論 5	435	出口 5	1591
世論 6	511	出口 6	1901
世論 7	407	出口 7	2059
世論 8	396	出口 8	2681
世論 9	617	合計	20860
世論 10	516		
世論 11	795		
合計	5803		

よく使っている質問である「投票の基準（政策、政党、人柄、経歴、無党派色、その他）」を予測に使用する共変量として採用することにし、2003年以降に行なわれた11回分の知事選挙での情勢調査データと、出口調査は6回分の知事選挙のデータを使用した。表1には、調査ごとの回収標本サイズを示した。以下の分析で用いる総標本サイズは、情勢調査については11調査あわせて5,803、出口調査は6調査20,860である。また情勢・出口の両調査において、支持政党や投票の基準の質問には、無回答も存在したが、これらは「無回答というカテゴリ」として処理した。

図2は、分析に用いるデータの情勢調査の結果と実際の開票結果を、図3は出口調査の結果と実際の開票結果をプロットしたものである。出口調査の結果は45度の直線付近にほとんどの点が集中しており、出口調査の結果は開票結果にかなり近いことがわかる。これと比べると、情勢調査の結果は45度の直線付近のばらつきが大きい。今回使用したデータでは情勢調査のサンプルサイズは出口調査と比べて小さく、そのため標本誤差が大きくなっているが、投票日

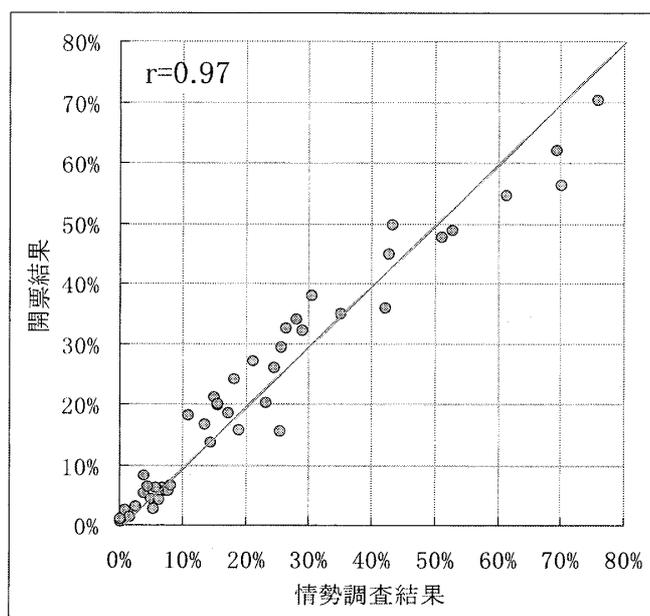


図2 情勢調査の結果と開票結果

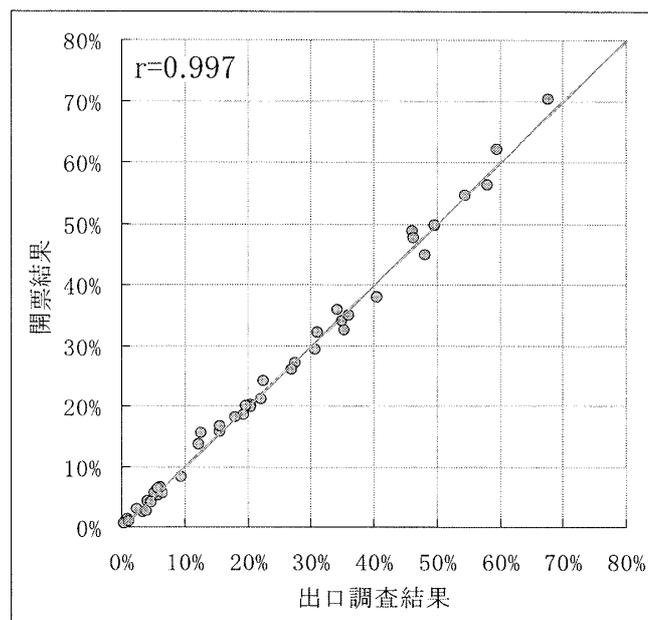


図3 出口調査の結果と開票結果

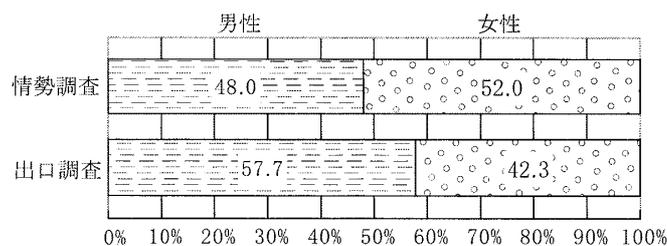


図4 調査ごとの男女比の比較

一週間前の情勢調査では、「新人候補の名前が挙がりにくい（実際の得票率は増加する）、著名人候補の名前が挙がりやすい（実際の得票率は減少する）」など独特の偏りをもっている。情勢調査結果のばらつきにはそのような偏りも含まれていると考えられる。

#### 4.2 両調査の回答者の特性

次に、情勢・出口両調査の回答者の特性を比較する。図4は各調査の男女構成比を表している。情勢調査では女性がやや多く、逆に出口調査は6割弱が男性の回答者となっている。しかし実際の投票結果を見ると、女性が男性をやや上回っている選挙が多い。出口調査で男性の回答者が多いのは、対象となった女性に協力をお願いしても、一緒にいるその夫の方が回答しようとして女性からの回答を得られない、というケースがあることも影響しているだろう。投票行動が性別によって大きく異なる選挙がある場合は注意が必要であるが、ここではデータをそのまま使用した。

図5は年齢の構成比を表している。情勢調査の回答者は、調査が実施される週末の低い在宅率や固定電話の保有率低下などの影響を受けて20歳代などの若い層の比率が低くなっている。また、図6は、各調査の「候補者を選ぶ際、何を基準にしますか（情勢調査）」「あなたは何を基準に投票しましたか（出口調査）」という質問に対する回答比率を表したものである。注目されるのは、情勢調査で「政策」を基準に投票すると回答している人が4割を超えているのに対し、出口調査ではその約半分となっていることである。これは投票に行くつもりはないが、

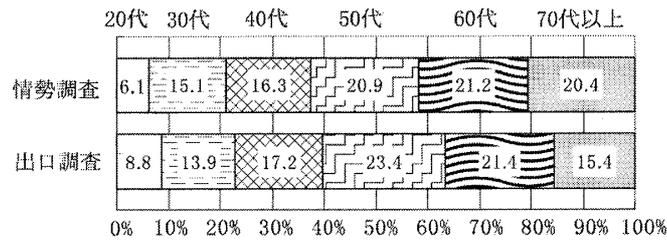


図5 調査ごとの年齢構成の比較

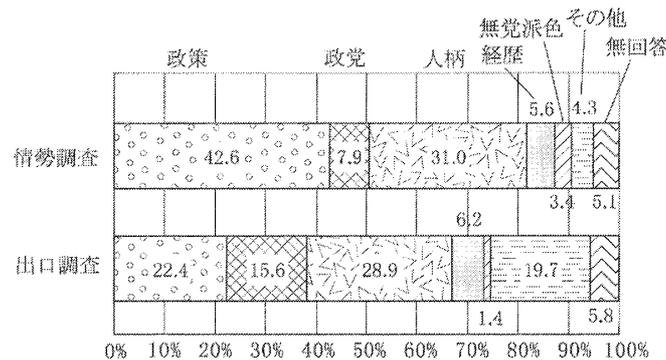


図6 調査ごとの投票の基準の比較

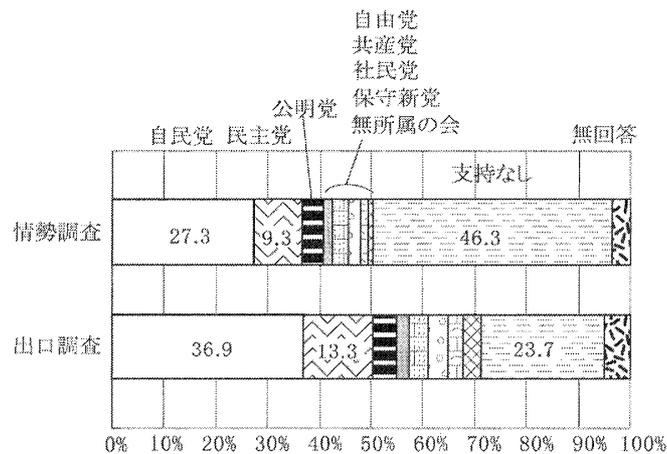


図7 調査ごとの支持政党の比較

たまたま情勢調査の対象者となってしまったので建前的に「政策で選ぶ」と答えている人が多いのではないかと推察される。そして、図7は支持政党に対する回答の比較である。ここで目立っているのは、出口調査では「支持する政党がない」と回答した人が情勢調査と比べて少ないことである。これは情勢調査で「支持政党はない」と回答した人の中には政治に無関心な層が多く、この無関心層があまり投票に行かなかったために、出口調査で抽出されなかったのだと考えられる。

#### 4.3 傾向スコアの推定と共変量のバランス

傾向スコアの推定には一般的に使われているロジスティック回帰分析を用いた。共変量は、4.1節で挙げた全変数（性別、年齢、投票区の市群別、支持政党、投票の基準）に加え、変数同士の積をいくつか投入した。図8は調査ごとの推定された傾向スコアの度数分布を表している。情勢調査の回答者群では0.3付近がピークとなっており、その一方出口調査の回答者群で

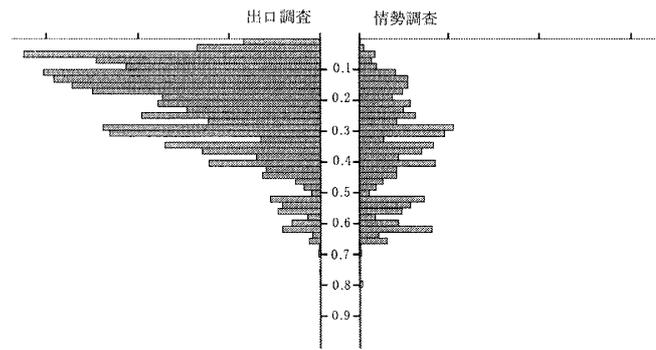


図8 傾向スコアのヒストグラム

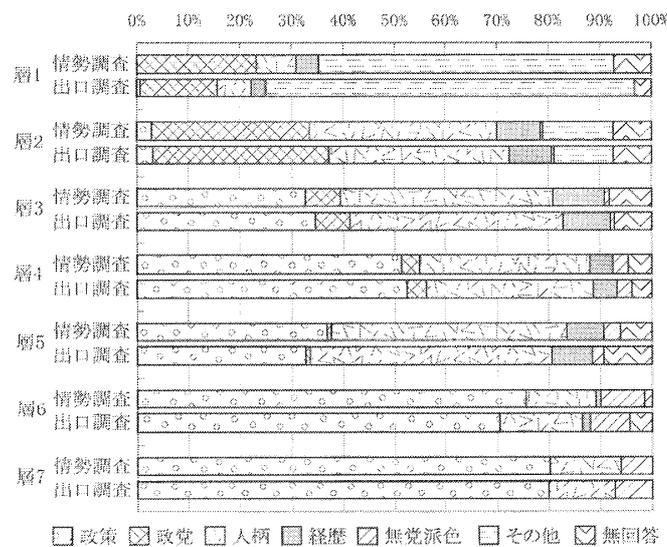


図9 傾向スコアの層ごとにみた「投票の基準」

は0.1付近がピークとなっている。そして、両調査の回答者群とも0.7を超えたものはほとんどない。両方の調査対象者において $\hat{e}(\mathbf{x})=0.6$ 付近で小さな山があるが、これは「政策」を基準に投票する、支持政党を持たない30歳代と50歳代の女性という回答パターンをもつ集団が分布している。

ここで推定された傾向スコア $\hat{e}(\mathbf{x})$ によって共変量 $\mathbf{x}$ のバランスがとれているか、つまり $\mathbf{x} \perp z | \hat{e}(\mathbf{x})$ が成立するかについて確かめる。図9は情勢調査の回答者と出口調査の回答者を、推定された傾向スコア $\hat{e}(\mathbf{x})$  ( $0 < \hat{e}(\mathbf{x}) < 1$ )で10個の層(0.1刻み)に層別し、「投票の基準」の回答分布を表したものである。サンプルが少ない $\hat{e}(\mathbf{x}) > 0.7$ の層は省略した。情勢調査の回答者層全体と出口調査の回答者層全体では投票の基準の回答分布は大きな差があったが、傾向スコアの推定値で層別すると両調査の回答分布が近づき、バランスがとれていることが分かる。よって傾向スコアが同程度なら属性や意識が似ているということであり、傾向スコアによる調整の効果が期待できる。

#### 4.4 結 果

傾向スコアの推定に使用した情勢調査データ11選挙のうち、出口調査を行っていない1選挙を除いた10選挙の候補45人について補正方法A, B, Cを適用した結果を表2, 3に示した。表2は、45人すべての候補の結果で、表3は、その中から得票率が10%を超えた候補28人を抜き出したものである。これは低い得票率の候補は当落に絡む可能性は低く、予測にとっ

てあまり重要ではないからである。なお方法 A, B では, 図 8 を見ると, 情勢調査の回答者には傾向スコアが 0.1 付近でのサンプルが少ないため  $\bar{y}_1^{(1)}$  の値が不安定になるので傾向スコアが 0 から 0.2 をひとつのグループとした。また, 0.8 以上のものは無いので,  $\bar{y}_1^{(h)}$ ,  $h=(1, 2), 3, 4, 5, 6, 7$  の加重平均を使った。

$$\begin{aligned} \bar{y}_1^{(A)} = & 0.53 \cdot \bar{y}_1^{(1,2)} + 0.21 \cdot \bar{y}_1^{(3)} + 0.15 \cdot \bar{y}_1^{(4)} \\ & + 0.05 \cdot \bar{y}_1^{(5)} + 0.04 \cdot \bar{y}_1^{(6)} + 0.02 \cdot \bar{y}_1^{(7)} \end{aligned} \quad (11)$$

を使用した。補正 B では変数選択の結果,  $\bar{y}_1^{(1,2)}$ ,  $\bar{y}_1^{(6)}$ ,  $\bar{y}_1^{(7)}$  が採用され,

$$\bar{y}_1^{(B)} = 0.0149 + 0.53 \cdot \bar{y}_1^{(1,2)} + 0.29 \cdot \bar{y}_1^{(6)} + 0.11 \cdot \bar{y}_1^{(7)} \quad (12)$$

を使った。

全体の候補者での結果 (表 2) を見ると, 補正 B (修正加重平均) を用いた結果が最も良く, 45 人の候補者のうち約 6 割の 27 人の予測値が出口調査の結果に近づき, それに伴って実際の開票結果にも近づいた。平均二乗誤差 (MSE) も補正 B によって最も小さくなっていることが分かる。補正 C は, 近づいた候補者数では補正 B の次に良かったが, MSE では補正 A より離れた。有力候補に限定した 28 人で見ても, 補正 B が近づいた人数, MSE とともに良く, A, C ではあまり改善が見られなかった。

図 10 は 45 人の予測値の補正前の点 (\*印) と補正後 (B) の点 (○印) をプロットしたものである。同じ高さの \*印と○印が同じ候補者を表しているが, \*印より○印のほうが 45 度の直線に近いものが改善された予測値である。

## 5. 最後 に

以上の結果を見ると, 補正によって改善された予測値は全候補の 6 割にとどまった。結果の良かった補正方法 B については, 新しいデータへの適合性の問題も残り, 今後検証していく

表 2 全候補 45 人の結果

	補正なし	補正 A	補正 B	補正 C
情勢調査結果が出口調査結果に近づいた候補者数		23/45(51%)	27/45(60%)	25/45(56%)
情勢調査結果が開票結果に近づいた候補者数		22/45(49%)	28/45(62%)	22/45(49%)
MSE (出口調査)	0.28%	0.31%	0.17%	0.38%
MSE (開票結果)	0.22%	0.23%	0.13%	0.29%

表 3 得票率 10 %を超えた候補 28 人の結果

	補正なし	補正 A	補正 B	補正 C
情勢調査結果が出口調査結果に近づいた候補者数		13/28(46%)	22/28(79%)	14/28(50%)
情勢調査結果が開票結果に近づいた候補者数		13/28(46%)	21/28(75%)	13/28(46%)
MSE (出口調査)	0.43%	0.49%	0.23%	0.59%
MSE (開票結果)	0.33%	0.36%	0.17%	0.45%

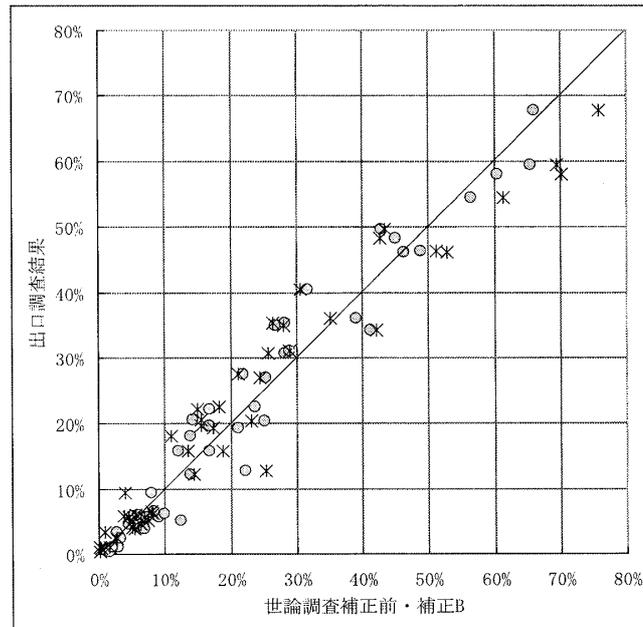


図10 補正前と補正Bのプロット

必要がある。しかし、情勢調査のデータを「どういう人が投票に行きそうか」という観点から補正を行うこの方法は、有効である可能性がある。

今回使用した補正前の情勢調査結果において、1位と予測した候補10人のうち1人が落選し、この当落判定を補正によって改善することは出来なかった。これだけでは補正によって当落判定が改善するかどうかは評価できないが、次点との差がどのくらいあるか、など得票率を予測することも重要であり、予測得票率をうまく補正できれば、それによって当落判定が改善される可能性は大きい。

ここで述べた手法をより本格的に使えるものにするには、改良を加える必要がある。傾向スコアによる調整は、Couper (2000) も指摘している通り、共変量  $\mathbf{x}$  の選択が手法の成否の鍵を握っている。共変量に関しては強く無視できる割り当て条件、つまり式(3)が前提になっており、これが成立しなければこの手法は無意味になってしまう。これは同じ対象者の1週間前の投票意志(情勢調査の回答)と実際の投票(出口調査の回答)を知らなければならず、直接は確認できない。しかし、これは同一パネルによる継続調査で選挙前と選挙後の調査を実施することによって  $y_0, y_1$  両方の観測が可能となる。このような追跡型調査を繰り返し行い、式(3)において、できるだけ独立になるような共変量を見つけだしていくことでこの手法が大きく改善される可能性がある。

その一方で、共変量の数を無制限に増やすことは、コストや出口調査の質問数の制限もあり簡単ではない。共変量の工夫と合わせて共変量の少なさを補うような手法も必要となる。

また、ここで使った分析では、情勢調査の回答者層で  $\hat{e}_0(\mathbf{x})=0.1$  付近のサンプルサイズが十分とは言えない。傾向スコアが0.1である集団が出口調査の回答者層の多くを占めていることを考えると、傾向スコアのオーバーラップ部分のサンプルを増やす工夫をすることによって補正がより安定するだろう。

そして、国政選挙では、候補者の政党色が強い、直前に支持なし層が減少するなどの特徴があり、地方選挙でうまくいく共変量そのまま適用できるとは限らない。そのため、それぞれの選挙の特徴を織り込んでモデルを微調整することが必要となってくる。また共変量として使

用している支持政党に関しても政党の合併や分離の可能性が常についてまわり、それまで使っていた予測モデルを作り直す必要がある。以上のような問題を1つずつ解決することによって、この手法はより実用的で役立つものとなるだろう。

さらに何より大切なのは、分析に使用するデータをていねいに、正確に集めることである。ここで議論した分析を有効に使うために、今後も信頼のおける調査ができるよう努めたい。

## 謝辞

本研究を進めるにあたり、数々の貴重なアドバイスを頂きました杉山高一先生、藤越康祝先生（中央大学）、並びに2人の匿名の査読者に深く御礼を申し上げます。

## 参考文献

- [1] Couper, M. P. (2000). Web Surveys: A review of issues and approaches. *Public Opinion Quarterly*, **64**, 464-494.
- [2] Hirano, K., Imbens, G., & Ridder, G. (2003). Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score. *Econometrica*, **71**, 1161-1189.
- [3] 長江一平 (2002). 毎日新聞の第19回参議院選電話調査. 行動計量学, **29-1**, 70-80.
- [4] 西平重喜 (1984). 統計調査法改訂版, 新数学シリーズ8, 培風館.
- [5] 福田昌史 (2005). 調査データの補正について, 2005年度統計関連学会報告集.
- [6] 星野崇宏 (2003). 調査データに対する傾向スコアの適用. 品質, **33-3**, 44-51.
- [7] 星野崇宏・繁樹算男 (2004). 傾向スコア解析法による因果効果の推定と調査データの調整について. 行動計量学, **31-1**, 43-61.
- [8] 松田映二 (2002). 朝日新聞社のRDD調査について. 行動計量学, **29-1**, 81-89.
- [9] Rosenbaum, P.R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika* **70**, 41-55
- [10] Taylor, H. (2000). Does Internet research work?, *Journal of the Market Research Society*, **42-1**, 51-63.
- [11] Taylor, H., Bremer, J., Overmeyer, C., Siegel, J. W., and Terhanian, G. (2001). The record of internet-based opinion polls in predicting the results of 72 races in the November 2000 US elections. *International Journal of Market Research*, **43-2**, 127-135.
- [12] 窪田知久 (2002). 読売新聞社の選挙情勢調査. 行動計量学, **29-1**, 62-69.