

# Topic 1

## 選挙における調査と予測報道

鈴木督久

2009年8月の衆院・総選挙は、民主党が300議席を超す圧勝をして政権交代という歴史的結果に終わった。2010年7月には参院選が実施され、民主党政権の1年間に対する有権者の評価が下される。選挙があると報道機関は有権者に対する情報提供として、選挙予測報道をする。選挙予測は正確さを要求され、それが報道機関の社会的信頼を支えている側面もある。

報道機関の選挙予測は主に新聞紙面とテレビ番組で報道される。新聞社は投票日前に情勢報道をし、テレビ局は投票当日に開票速報を放映する。この2種類の選挙予測を以下のように区別しておこう。

- (1) 公示後から投票日まで報道する事前の予測（主に新聞紙面）
- (2) 投票後から開票終了までの当日報道する予測（主にテレビ番組）

どちらの予測報道も量的調査と質的調査を実施した結果を利用している。量的調査とは標本調査であり形式的には世論調査と同じだが、当日の開票速報で利用している標本調査は特に出口調査と呼ばれている。出口調査の実施主体はテレビ局だけでなく新聞社・通信社でも実施しており、開票当日は新聞社のWEBサイトで開票速報に使われたあと、翌日の紙面において有権者の投票結果の行動分析記事などに利用されている。質的調査とは記者による取材調査のこ

とである。主な内容は選挙区の関係者・支持組織などに対する取材、有権者の反応などの現地観察、さらに開票所における開票結果の取材も含むあらゆる情報収集活動の総称である。

### 事前の情勢予測と報道

2009年8月の総選挙で全国紙・通信社は表に示したように、民主党が300議席超で圧勝して政権交代すると予測報道し、実際にそうなった。報道機関の選挙予測の最大目的は政党別獲得議席数であり、一面トップで報道される。これは選挙後の政治体制を決めるからであり、小選挙区制の下ではどの政党が過半数の議席を獲得するかが焦点となる。第2の目的は各候補者の個別の当落である。これは新聞の中間に選挙区別の情勢解説記事として掲載される。

標本調査の測定方法としては、衆院に小選挙区比例代表並立制が導入された1996年以降は電話法が主流となった。標本抽出方法としてはRDD（ランダム・デジット・ダイヤリング）法が採用されている。RDDサンプリングは固定電話世帯を枠母集団として無作為抽出した番号に電話し（第一段）、世帯内で調査対象者1人を無作為抽出して質問する（第二段）。質問内容は小選挙区と比例代表への投票先などを中心に10項目程度である。回収標本サイズは各選挙区で数百人前後だが、衆院の場合は300選挙



すずき とくひさ

株式会社日経リサーチ取締役。

東京大学大学院人文社会系研究科，早稲田大学政治経済学部非常勤講師。

1982年早稲田大学第一文学部卒業。同年日経リサーチ入社。90年日本経済新聞社政治部記者。世論調査室長，マーケティング局長などを経て，05年より現職。

主な著書は、『入門 共分散構造分析の実際』（共著，講談社），『多変量解析実例ハンドブック』（分担執筆，朝倉書店）など。

区もあるので調査全体の計画標本サイズは数十万人という大規模な調査プロジェクトとなる（表の標本サイズ欄参照）。

調査における投票意向，つまり各候補者に関する標本支持率は実際の得票率と同じではない。そこには標本誤差のほかに非標本誤差も含まれている。予測システムの第一段階は標本支持率から非標本誤差をできるだけ除去する「補正」という作業が実施される。補正結果を標本支持率とは区別して推定得票率と呼ぼう。次に標本誤差を考慮して推定得票率から各候補者の当選確率を求める。最後に当選確率を積み上げた結果としての政党別獲得議席数を区間推定する。比例代表は全国11に分類されたブロックごとに小選挙区とほぼ同様の方法で予測分析して政党別の獲得議席数の区間推定をする。これらのデ

ータ解析・予測分析は調査のあと数時間で終了しなければならない。大勢のスタッフが周到な準備をして進めるが，どれだけ準備しても現場はいつでも「絶対に失敗できない」という強い緊張感と時間制約の中で仕事をすることになる。

この推定結果をそのまま紙面に反映する新聞社もあるが，小選挙区に関しては記者を集めて取材情報を加えた検討をしたのち，統計的推測の結果を修正して紙面化する社もある。以上のような予測プロセスの結果として選挙前の新聞紙上に「民主が300議席超の勢い」という見出しが出るのである。

## 開票日の予測と報道

テレビ局は投票が終了する午後8時をめぐりに番組をスタートし，冒頭から当選確実（当確）の候補者を次々に発表していく。「なぜ開票率0%なのに当確なのか」「なぜ開票数が下回っ

ている候補者に当確が出るのか」という素朴な疑問をもつ視聴者もいる。

これは出口調査などの予測結果を報道しているからであり，開票結果の伝達ではないというのが疑問への回答である。調査と取材による質量両面の予測分析の結果から，開票前に当確と断定できる候補者が存在するということである。一方で接戦の小選挙区に関しては，開票結果と照合しながら慎重に判断していく。テレビ局でも「当確のミスは絶対に許されない」という緊張感に支配され，多数のスタッフが選挙報

表 新聞社・通信社による事前の選挙情勢予測調査・報道と投票結果

調査主体	掲載	報道紙面における見出し	議席数の予測値		標本サイズ
			民主	自民	
朝日(1)	8/20	民主、300議席うかがう勢い 自民苦戦、半減か	G	G	60,277
読売(1)	8/21	民主、300議席超す勢い 自民激減 公明は苦戦	G	G	109,893
日経	8/21	民主 圧勝の勢い 300議席超か 当選圏 自民半減以下も	G	G	109,893
毎日	8/22	民主320議席超す勢い 自民100議席割れも	318-330	68-108	77,858
共同	8/23	民主、300議席超す勢い 自民は100前後か	G	G	155,148
朝日(2)	8/27	民主、320議席獲得も 自民激減 100前後	307-330 321	89-115 103	130,879
読売(2)	8/28	民主 圧勝の勢い保つ 自民 激戦区で猛追	G	G	85,777
投開票	8/30	選挙結果の獲得議席数	308	119	

(注1) 朝日(1)の調査選挙区数は150。読売(2)は200。ほかは300選挙区が調査対象。

(注2) 予測値欄にGとあるのは，予測議席数の区間・点推定値を発表せずグラフなどで表示していることを示す。

道に動員されている。すべて開票される最後まで判定を保留する選挙区もある。こうして投票終了から数時間の選挙番組は、480議席が次第に決まっていくような流れで構成されるのである。

出口調査は投票所の出口で、調査員が投票者に面接して測定する。投票先の質問は調査員に見えないように自記式で封入回収する社もある。目標母集団は投票者だが、番組を午後8時にスタートするために調査を6時に終了して集計・分析することが多い。この場合は期日前投票者と午後6時以降の投票者の部分がカバレッジ誤差となる。この影響の構造と程度については各社で研究されており予測分析で考慮されるが、最近では増加する期日前投票に対応して、期日前投票所での出口調査も実施されている。

標本抽出は層化二段無作為抽出法で、地域的特性（地盤や産業・歴史背景など）を考慮して層化する。第一次抽出単位を投票区として確率比例抽出する。第二次抽出単位は投票者で、一定数を無作為抽出する。系統抽出法の利用が多いが、共同通信社では男女交互に抽出するなど各社で若干の実践的変法がある。抽出間隔は投票区内の投票者数が未知なので過去の実績などから推定して概算する。標本サイズは情勢調査と同程度かそれ以上に大規模である。

出口調査で得る標本支持率は事前の情勢調査よりも精度が高いが、やはり回答拒否もあり非標本誤差を含むので情勢調査と同様に標本支持率から推定得票率、当選確率を求めて政党別獲得議席数を区間推定する。午後6時の調査終了後ただちに分析を終え、2時間後の放送前には政党別の獲得議席数予測は終了しており、番組冒頭で「政権交代の模様だ」などと全体状況が報道される。午後8時以降は全国の投票所に配備された取材陣から本社に送られ続ける開票データを累積・集計して予測分析を進める。

テレビ局の予測議席数は番組の中で述べるだけで出版物はないが、2009年8月の総選挙においては、NHK（日本放送協会）の午後8時発表（以下同じ）は「民主298～329、自民84～131」という区間推定であった。ANN（テレビ

朝日）は「民主315、自民106」という点推定で、NHKの中央値「民主313.5、自民107.5」とほぼ同じだった。他の民放各社はいずれも点推定「民主320強、自民90後半」の近傍で共通していた。NHKやANNと比較すると他局は民主を過大（自民を過小）予測しており、新聞各社の事前情勢予測の値（表）に近かった。

ところで「開票率0%で当確」の仕組みであるが、300小選挙区における当選者と2位との得票率差の分布をみると、予測分析の現場実感を少しもてるだろう。図に得票率差を幹葉図で表示したが、集計すると20P（パーセント・ポイント）差以上の小選挙区が92あり、10P以上と10P未満がちょうど104ずつで、300選挙区がおおよそ100選挙区ずつ3等分されている。このように選挙は楽勝から接戦まで分布している。本当は標本支持率が推定得票率の分布を示したいところだが、それは公表されていないので得票率の分布から想像してほしい。標本支持率においても得票率と同じように楽勝から接戦まで分布している。得票率差には下限0という制約があるので分布の形状は少し異なるが、おおよそ近似した平均値を中心に分布していることは同

幹葉	#箱ヒゲ図
46 135	3 0
44 0	1
42 71	2
40 34	2
38 94	2
36 06	2
34 1405	4
32 2771226	7
30 19	2
28 004667845688	12
26 0366711334457	13
24 001124466901344	15
22 01128800134589	14 +-----+
20 01345779588	12
18 0122560012334579	16
16 01234473930011224456668899	26
14 000226779012344678	18 *--+-*
12 0234456677778800134409	22
10 001123345780000333456679	24
8 00133579011135777889	20
6 0000234466889023577789	21 +-----+
4 1111356778001144568	19
2 23445133455679999	17
0 22456667730001222334488889	25
-----+	

図 小選挙区の当選者と次点との得票率差分布（2009年総選挙）。  
（注）栃木3区（渡辺喜美が95.3%で当選）は外れ値として除いた。

じである。

実態はポイント差だけで判定するのではなく、当選確率も計算し取材情報も加味しているのだが、大雑把な状況を説明することがここでの目的である。たとえば「出口調査で20P以上の差をつけた候補者が落選した事例は過去にない」というような経験的事実を報道機関の予測担当者はもっていて、そのような圧倒的に優勢の候補者が開票率0%で当確を打つ候補となる。10P以上の差で落選した例は皆無というわけではない。99%の信頼区間でよい、つまり「1選挙区くらい当確を間違えてよい」という判断はないので、あらゆる情報を総合しながら当確の間違いを避ける。開票情報も単に何割の開票が終了したかではなく、どの地域の票が集計されたのか、つまり地盤という質的判断なども加えながら当確を打ってよいかを判断する。5P未満というような大激戦の選挙区は52あり、もちろん調査データでも接戦状況を示している。このような接戦の選挙区は急いで判定はせずに、開票結果を待つことになる。

## 予測分析の方法

情勢調査でも出口調査でも標本支持率から推定得票率を求める方法は回帰分析モデルが一般的である。過去の選挙における標本支持率などを独立変数とし、得票率を従属変数とする。独立変数の選択とか、ロジスティック回帰モデルを使うとか、細部は各社で違いがあるが本質的には過去の調査データから構成したモデルに今回の調査データを代入した予測値が推定得票率である。推定得票率には標本誤差しか含まれていない状態にすることが目標で、これを「補正」と呼ぶことが多い。この用語は偏りを除去するという実感からきている。

回答を拒否する人は多い。回答してもそれはウソかもしれない。投票所で気が変わるかもしれない。選挙に行かないかもしれない。有権者は投票態度を決めている固定層から、投票所に行ってから考えるような浮動層まで分布している。候補者特性による影響も受けるだろう。あらゆる非標本誤差を含んで偏る。本当は得票率

を測定したい。標本支持率が得票率から偏りをもつほど構成概念妥当性が低いともいえる。しかし予測にとっては妥当性より信頼性が重要である。偏りは不可避である。そうであるなら積極的に偏りがあってもかまわないと思うことにする。偏りを恐れず忌避しない。偏りはいつも安定して同じように偏ることが、つまり測定の信頼性こそが重要なのである。そのためには調査管理が重要である。いつも同じやり方で調査を実施してデータを蓄積する継続性こそが基本的に重要である。分析より収集が重要だと言い切ってもいい。そうすれば偏りを構造化できる。それが補正の意味である。

こうして求めた推定得票率から、次に当選確率を求める。この計算方法としては頻度論的なアプローチが一般的である。判別分析モデルを利用する場合もある。候補者個人の当選確率と政党別の獲得議席の区間推定を同時に求めるにはブートストラップ法も利用される。ベイズの方法による当選確率の計算も適用できると思われるが、現状では採用している報道機関はないと思われる。

2009年8月の総選挙における選挙予測は、総括的に述べれば「民主が300議席超で政権交代」という情勢が最後まで揺るぎなかったという意味で成功したが、民主党が320議席という予測は過大であった。接戦の選挙区で自民党の実力者たちが文字通り実力を発揮した。その傾向は終盤の情勢調査や出口調査に出ていたが、ギリギリの攻防であり統計的な判定の範囲外であったとも、調査と予測モデルの限界であったともいえる。

## 文献

- 福田昌史 (2008). 出口調査の方法と課題 行動計量学, **35**(1), 59-71.
- 林知己夫・高倉節子 (1964). 予測に関する実証的研究——選挙予測の方法論—— 統計数理研究所彙報, **12**(1), 9-86.
- 稲井田茂 (2009). 午後8時1分に「当確」を打つ 出口調査の手法 *Journalism*, **23**(1), 34-43.
- 仁平俊夫 (1996). 選挙と出口調査——NHKの手法と課題—— 行動計量学, **23**(1), 20-27.