

# 日本の世論調査における妥当性と信頼性に関する研究

## —統計分析を用いた比較検討—

大山 康平\*

### 【要旨】

一般市民の政治的意見・スタンスなどを調査しているのが「世論調査」である。日々世論調査の結果が報じられ、調査主体によって結果が異なることもある。なぜ、マスコミによって結果が異なるのか。社会調査法から見て今の世論調査は、世間一般の政治的意見を適切に調査・分析したものであるといえるのであろうか。世論調査に対しては、ワーディング・サンプリング等様々な観点から批判・疑問の声が上がっている。しかしながら、マスコミ各社によって実施された過去の世論調査を統計的に分析・比較検討した先行研究は殆どなされていない。本研究では、入手可能なマスコミ各社によって実施された過去の世論調査結果のデータに統計解析を施し、マスコミ各社発表の世論調査の結果間に① 統計的有意差が見いだされるのか、② 統計的差があった場合には、その背景を検討することを研究目的とした。その結果、世論調査には統計的な差があり、その背景には重ね聞きや携帯電話を調査に含めるかなどの調査設計部分が影響していることが示唆される結果になった。

キーワード：世論調査 社会調査法 重ね聞き 多重比較法

## 1. 序論

近年マスコミが行う世論調査に対して、風当たりが強くなっている。その理由として、選挙情勢調査などを含む世論調査結果と実際の選挙の結果に乖離が生まれたことが考えられる。アメリカメディアは「隠れトランプ」を見抜けず、日本では維新の会支持層を世論調査ではとらえきれていないといった事象が起きた。このように近年、国内外で世論調査に対して批判や疑問点などが指摘されている。このような事例をもとに、本研究では、世論調査を取り巻く状況を整理したうえで、マスコミ各社の世論調査の集計データに統計解析を施し、世論調査の信頼性・妥当性を検討することを目的とする。

世論調査に対しては科学的な調査といえるのか、また標本抽出が正しくなされているのか、質問は適切かなど様々な点について批判がなされている。調査設計面での批判としては、世論調査で多い 2 択式の回答が、多様な考え方を表出させることを妨げ、極端な印象を引っ張っていると批判しているという批判（菅原 2016）や、現在の世論調査手法では社会システムとかけ離れたものになっているので計量社会学の知見を取り入れ、より社会と密接なものにしていくべきだと主張しているもの（吉川 2016）、メタ語用論の観点から

---

\* 関西学院大学大学院総合政策研究科博士課程前期課程 (efi97694@kwansei.ac.jp)

post-truth の問題の 1 つとして重ね聞きが数値を高い方へ導くこと、また選択肢の設定法で変わる部分が多いことから、主観的真実があたかも客観的真実のごとく報じられているのではないかと指摘しているもの（沖田 2018）、現行の世論調査の方法では、政党支持の強度を尋ねるものはほぼなく、自らの調査をもとに世論調査で捕らえ切れていない「完全中立層」の存在を把握したうえで支持政党の強度が、選挙の結果や投票行動に密に関連していると考察し、政党支持を図る際にはその強度も調べるべきであると指摘しているものがある（井田 2013）。

世論調査をどう受け止めるのか、またどのように活用すべきなのかという議論では世論調査を行うメディアは調査票を有権者と政治を結びつけるものになるような形にし、社会問題を解決するための方向性を探るものになるようにすべきであると主張し、そのような形にすれば世論調査批判は減ると述べるという主張や（菅原 2011）、世論調査が世論を客観的に測定するためにあるものであるとし、政治が直接左右されるものではなく、「考慮されるべき素材」になるべきと主張し、完璧な世論調査は不可能であるからこそ、どのように偏っているのかを明らかにし、その結果を共有することに意義があるという主張（盛山、2016）、世論調査、世論調査報道において「ふつうのひと」の声が取り上げられず、弱者やノイジーマイノリティにより注目がいく仕組みになってしまっていると指摘し、実際の世論とのかい離が生まれてしまっていると指摘する。また、世論調査で取り上げるトピック・結果が時に「外れ値」になる危険も指摘しているもの（佐藤 2016）がある。

また固定電話中心、週末主体が基本の世論調査に対しての批判に対しても研究がなされている。それらでは携帯電話を含めることの必要性が論じられているが、問題点も指摘されている。また、週末に外出する層をとらえきれない点についてはその層の存在を認めつつも、結果には影響がないという先行研究がなされている（藤木 2014）。

## 2. 方法

本研究では大学図書館のデータベースや一般公開されている世論調査の集計データを利用し収集した。調査対象は、NHK、朝日新聞、毎日新聞、産経新聞、読売新聞、日経新聞、JNN の 7 社とし、分析の対象期間は 2014 年 1 月から 2018 年 6 月までの 4 年半分とした。そのデータに統計解析を施して世論調査の実態について検証、比較検討を行うことを目的とする。

分析に当たっては主要マスコミの世論調査集計結果を利用し、差異法などを変数として比較検討を行うこととした。その際、日経は重ね聞きしているものとしていないものを両方公表しているが、本研究では両方をそれぞれの結果としてとして扱うこととする。分析に当たっては統計ソフト SPSS 25 を使い、独立したサンプルの t 検定、一元配置の分散分析とその後の検定として Bonferroni の多重比較などを行い、統計的仮説検定を行った。また各社ごとに少しずつ異なる調査設計を変数化し、設計が結果にもたらす影響を検証した。また「統計的な比較のできない項目」として「米軍普天間飛行場の名護市辺野古への移設問題」を事例とし、ワーディングによる世論調査を利用した世論誘導の考察も併せて行う。

### 3. 結果

#### 3.1 社ごとの比較

まず初めにすべてのデータを対象に一元配置の分散分析とその後の検定として Bonferroni の多重比較を行った。分析の結果有意水準 5% で統計的な差があることが認められた。また分類化も予見される結果となった。

多重比較の結果は表 2 のようになり、各社の対応関係は以下ようになった

朝日・毎日 < NHK・日経 < FNN・日経〈重ね聞き〉・読売 < JNN

表 1 内閣支持率とマスコミ各社の分散分析の結果

内閣支持率の平均値と標準偏差		F 値 (自由度)
NHK	平均値=47.76 SD=5.6833	26.125 (7) ***
朝日	平均値=43.52 SD=5.4070	
毎日	平均値=43.84 SD=6.8621	
産経/FNN	平均値=50.32 SD=5.6889	
読売	平均値=53.02 SD=6.7446	
日経	平均値=45.90 SD=5.9203	
日経〈重ね聞き〉	平均値=51.51 SD=6.6944	
JNN	平均値=55.39 SD=5.7085	

\*\*\*p<.001 水準で有意

図 1 マスコミ各社の内閣支持率の平均値のグラフ：標準誤差表

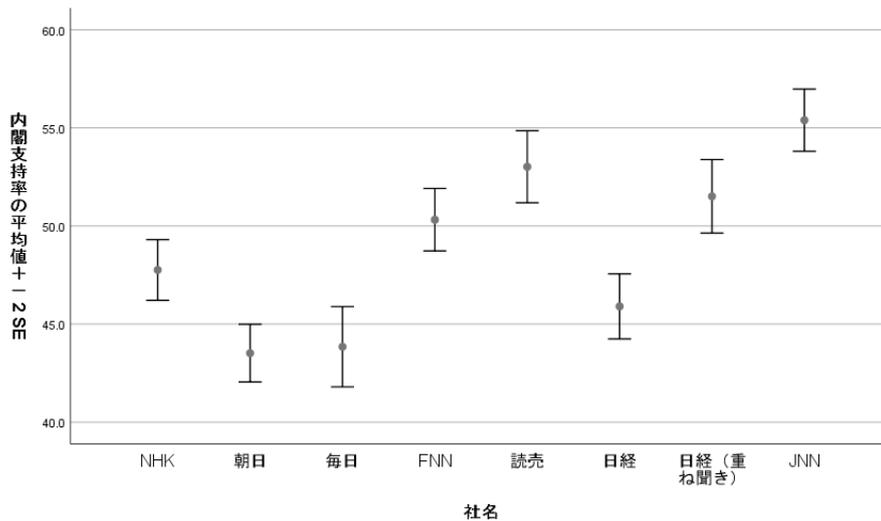


表2 多重比較の結果要約表（内閣支持率）

多重比較で差のあったもの：平均値の差（絶対値）と標準誤差			
NHK	朝日	平均値の差=4.2407*	SE=1.1734
	毎日	平均値の差=3.9148*	SE=1.2306
	読売	平均値の差=5.2593*	SE=1.1734
	日経〈重ね聞き〉	平均値の差=3.7505*	SE=1.1905
	JNN	平均値の差=7.6330*	SE=1.1846
朝日	FNN	平均値の差=6.8011*	SE=1.1905
	読売	平均値の差=9.5000*	SE=1.1734
	日経〈重ね聞き〉	平均値の差=7.9913*	SE=1.1905
	JNN	平均値の差=11.8738*	SE=1.1846
毎日	FNN	平均値の差=6.4752*	SE=1.2470
	読売	平均値の差=9.1741*	SE=1.2306
	日経〈重ね聞き〉	平均値の差=7.6654*	SE=1.2470
	JNN	平均値の差=11.5479*	SE=1.2413
FNN	日経	平均値の差=4.4176*	SE=1.2074
	JNN	平均値の差=5.0727*	SE=1.2015
読売	日経	平均値の差=7.1166*	SE=1.1905
日経	日経〈重ね聞き〉	平均値の差=5.6078*	SE=1.2074
	JNN	平均値の差=9.4903*	SE=1.2015
日経〈重ね聞き〉	JNN	平均値の差=3.8825*	SE=1.2015

\*p<.005 水準で有意 SE は標準誤差

### 3.2 〈重ね聞き〉についての分析

次に各社間の調査手法の違いの中から〈重ね聞き〉に着目し、独立したサンプルのt検定を行った。この際、他の要因の影響は排除せず、あくまでも一要因としての重ね聞きに着目する。分析の結果、内閣支持率 (t=-5.58)、自民党支持率 (t=-13.02)、共産党支持率 (t=-3.20)、支持政党なし (t=4.09) の平均値がいずれも 1%水準で統計的有意差があることが判明した。

表3 内閣支持率の〈重ね聞き〉の有無による比較：t検定の結果

内閣支持率の平均と標準偏差		t 値（自由度
〈重ね聞き〉あり	平均値=52.29 SD=6.73	-5.58 (410) ***
〈重ね聞き〉なし	平均値=47.85 SD=7.13	

\*\*\*p<.001 水準で有意

**表4 自民党支持率の〈重ね聞き〉の有無による比較：t検定の結果**

自民党支持率の平均と標準偏差	t 値 (自由度)
〈重ね聞き〉あり 平均値=35.17 SD=3.48	-13.02 (410) ***
〈重ね聞き〉なし 平均値=40.28 SD=3.41	

\*\*\*p&lt;.001 水準で有意

**表5 共産党支持率の〈重ね聞き〉の有無による比較：t検定の結果**

共産党支持率の平均と標準偏差	t 値 (自由度)
〈重ね聞き〉あり 平均値=3.84 SD=1.06	-3.20 (410) ***
〈重ね聞き〉なし 平均値=3.49 SD=0.95	

\*\*\*p&lt;.001 水準で有意

**表6 「支持政党なし」の〈重ね聞き〉の有無による比較：t検定の結果**

「支持政党なし」の平均と標準偏差	t 値 (自由度)
〈重ね聞き〉あり 平均値=37.48 SD=6.13	4.09 (409) ***
〈重ね聞き〉なし 平均値=40.56 SD=6.86	

\*\*\*p&lt;.001 水準で有意

### 3.3 「選択肢の分割」の影響

次に内閣支持率の選択肢が結果に与える影響を検証する。ここでも独立したサンプルのt検定を行う。その結果、4択を2択に集計しているJNNと2択を採用しているそれ以外の6社との間には、内閣支持率/内閣不支持率ともに統計的有意差があることが判明した。

**表7 内閣支持率の選択肢の数による比較：t検定の結果**

内閣支持率の平均と標準偏差	t 値 (自由度)
2 択 平均値=48.05 SD=7.02	-7.197 (410) ***
4 択 平均値=55.39 SD=5.71	

\*\*\*p&lt;.001 水準で有意

**表8 内閣不支持率の選択肢の数による比較：t検定の結果**

内閣不支持率の平均と標準偏差	t 値 (自由度)
2 択 平均値=36.04 SD=6.47	-7.266 (410) ***
4 択 平均値=42.92 SD=5.66	

\*\*\*p&lt;.001 水準で有意

### 3.4 比較検討できない「質問」について

ここまでマスコミの世論調査の質問項目のうち、「内閣支持率」「政党支持率」の項目を取り上げ、統計分析を行ってきた。世論調査で政策課題など尋ね方が社によって違うものは質問に意図がある可能性もあるためこれまで行ったような比較ができない。その一例が以下の高度プロフェッショナル制度に対しての各社の聞き方の違いである（各社の世論調査より抜粋）。

表9 世論調査における「働き方改革」の質問文例

社名	実施年月	文章
産経/FNN	2018/2	所得の高い一部専門職に限り、労働時間の規制対象から外す「高度プロフェッショナル制度」の導入に賛成ですか、反対ですか。
JNN	2018/3	「働き方改革関連法案」では、高収入の一部の専門職を労働時間の規制から外す「高度プロフェッショナル制度」を創設する方針ですが、野党は「残業代ゼロ法案だ」と批判しています。あなたはこの制度に賛成ですか、反対ですか。
朝日新聞	2018/3	政府は、専門職で年収の高い人を労働時間の規制の対象から外す「高度プロフェッショナル制度」を働き方改革関連法案に盛り込む方針です。高度プロフェッショナル制度を法案に盛り込むことに賛成ですか。

このように、各社によって大きく質問項目が違うものも多い。この場合、単位同質性<sup>1</sup>が確保できないため、統計分析を行うことは不可能であると考えられる。また、質的な比較もできないと思われる。

## 4. 考察

分析の結果から世論調査には統計的有意差があることが判明した。その差を生む要因としてはワーディングなどの調査設計部分にあるのではないかと考えられる結果になった。またマスコミ間によって<重ね聞き>を《行う/行わない》、大きく質問文が異なる、このような世論調査が行われている現状から、果たして我々が日ごろ接している世論調査は信用できるものなのかという疑問が出てくる。しかし重ね聞きや細かいワーディングなどの差はあるものの、先行研究でも議論されているように、内閣支持率や政党支持率などのサンプルかつ同じ質問を繰り返しして調査している項目についてはマスコミ間に差はあれども調査として信頼性のあるものであると考えられる。一方で政策 이슈を問うものに関しては、ワーディングに各社のイデオロギーなどが出ているものもあり、また同じ質問を

<sup>1</sup> 社会科学における比較で同じ意味の値ということ。二つの観察単位における従属変数の期待値が同じであるならば、この二つの観察単位は同質的であると考えられるというもの。

繰り返し行っていない場合もあり、比較が困難であり、調査としても不十分なものであると考えられる。このような調査に基づく報道は不誠実・不正確なものであると言わざるを得ない。

## 5. 結論

結論として世論調査には実施主体間で統計的有意差があり、ワーディングやサンプリングの違い、重ね聞きなどマスコミ各社による設計の違いがその差を生むことが明らかになった。この結果から考えられるのはマスコミ間統一ルールの作成が必要であると考えられる。また誘導的・恣意的と捉えかねられないような質問文での調査を行わない、同じ質問を最低 10 回以上の調査で継続して行うなど社会調査として自然な形での調査を行うことが求められていると考える。

本稿で主に取り扱った〈重ね聞き〉については、調査において〈重ね聞き〉した場合は必ず二種公表〈重ね聞き〉(する前/した後)をすべきであると考えられる。本研究では、〈重ね聞き〉することが回答の上方修正を招くことが明らかになった。そうである以上、一般市民に提供すべき情報は都合よく修正したものではなく調査結果の実態そのものである。〈重ね聞き〉によってより支持/不支持がわかるという点から重ね聞きそのものについては否定しない。しかしもたらす効果を合わせて公表することは必須であると考えられる。選択肢の分割についても同様に 4 択のまま公表すべきであると考えられる。

このような統一ルールの作成、あるいは同じ質問を繰り返し調査するといった科学的手法を用いた調査を行うことによって、世論調査がいたずらに政治状況をあおっているとの批判も解消されるとともに、世論調査の本来の目的である、「その時点における人々の政治的スタンスなどを測る」につながると考える。世論調査がポピュリズムにつながりかねない現況は非常に危険な状態であると考えられる。マスコミ各社は世論調査の意義を再度深く考え、政治と社会をつなぐ貴重な物差しであることを認識し、多面的で専門性を保ちながら、専門知識を持たない人であってもその実態がわかるようなものであるべきであると筆者は考える。

## 謝辞

本研究に際して時に厳しく、時にやさしく常に的確なご指導をいただきました李政元教授に厚く御礼申し上げます。また途中で病気になった際には励まし、またジャーナリストとしてのアドバイスを多く頂戴した小池洋次元教授（現関西学院大学フェロー）にも感謝申し上げます。また、持病が悪化した際に、親身になって支えてくださった総合支援センター学生相談支援室の皆様、そして総合政策学部職員の皆さまにも深く感謝致します。

**【参考文献】 一部のみ。全ての参考文献については修士論文を参考にされたい**

井田正道 (2013)『世論調査を読む—Q&A から見る日本人の“意識”』 明治大学出版会.

岩本裕 (2015)『世論調査とは何だろうか』 岩波新書.

菅原琢 (2016)『政治と社会を繋がないマス・メディアの世論調査 (特集 世論をめぐる困難)—(世論測定をめぐる困難)』, 放送メディア研究= *Studies of broadcasting and media*, (13), 57-78.

盛山和夫 (2016)『世論調査と政治的公共空間』 (特集 世論をめぐる困難)—(世論測定をめぐる困難), 放送メディア研究= *Studies of broadcasting and media*, (13), 43-56.

鄭躍軍 (2015)『日本における世論調査の信頼性と妥当性の統計学的検証 (第 3 部 研究発表, < 特集> 特別研究大会報告)』, 日本世論調査協会報「よろん」, 115, 46-50.

吉川徹 (2016)『社会のしくみと世論の距離』 (特集 世論をめぐる困難)—(世論が生まれる社会の困難), 放送メディア研究= *Studies of broadcasting and media*, (13), 215-230.

吉田徹編著他 (2018)『民意のはかり方「世論調査×民主主義」を考える』 法律文化社.