

## 回収率にみる調査の病状

－日本とアメリカの現況－

Limitations of Surveys Evaluated by Response Rate:  
Current Cases in Japan and the U.S.A.



松田 映二

1. はじめに
2. 各調査手法の回収率
  - 2-1. 面接調査の回収率
  - 2-2. 電話調査の回収率
  - 2-3. 郵送調査の回収率
3. アメリカの調査状況
4. 終わりに

### 〈要旨〉

回収率は、調査の質を見極めるための最重要指標である。本稿では、日本の各調査手法による回収率のトレンドを分析し、調査の病状を診断する。面接調査の回収率は、報道機関では低下傾向にあるにもかかわらず、内閣府では回復傾向にあるのは奇妙である。日本ではRDD調査の回収率を正確に計算しないため、調査の質の低下が判明しにくい。しかし、アメリカでは回収率を計算する6つの明確な定義がある。計算式によってアメリカのRDD調査の回収率低下は明確になっている。そのため、RDDに代わる新しい試みが進行している。

Response rate (RR) is the most important index for evaluating survey quality. This article analyzes RR trends of different survey methods to figure out their limitations. The RR of face-to-face surveys conducted by Japanese news media continues to decline. However, the RR of the Cabinet Office of the Japanese Government indicates recovery as a V-curved line. The difference between the two is unusual. As the RR of telephone surveys conducted in Japan is not calculated accurately, it is not possible to evaluate the quality of the survey results. The American Association for Public Opinion Research (AAPOR) provides six definitions of RR. Based on these definitions, the RR in random digit dialing (RDD) surveys in the U.S.A shows a rapid decline. Therefore, new experiments are being conducted in the U.S.A. to replace RDD.

## 1. はじめに

松田(2014)は、2004年3月4日に開催された埼玉大学社会調査研究センター設置記念シンポジウムにおいて、「伝統的な統計調査を捨てて新たな方法論に切り替えるパラダイムシフトは、調査への信頼を取り戻すために求められる本質ではない。代わりに安定的に80%以上の高回収率を獲得できる『目で見える』複合調査の開発を目指す」(p.3)ことを提唱した。しかし、その道のりは険しい。『世論調査年鑑』によれば、日本で実施された回収率80%以上の世論調査の数は、1979年をピークに減少を続け2012年は104件にとどまっている(図表1)。

この回収率の低下傾向は、面接、郵送、電話などの各手法で共通に見られるのだろうか。政府の公式統計や報道各社の公開情報をもとに、回収率のトレンドを分析し、どのような課題があるのかを明らかにする。

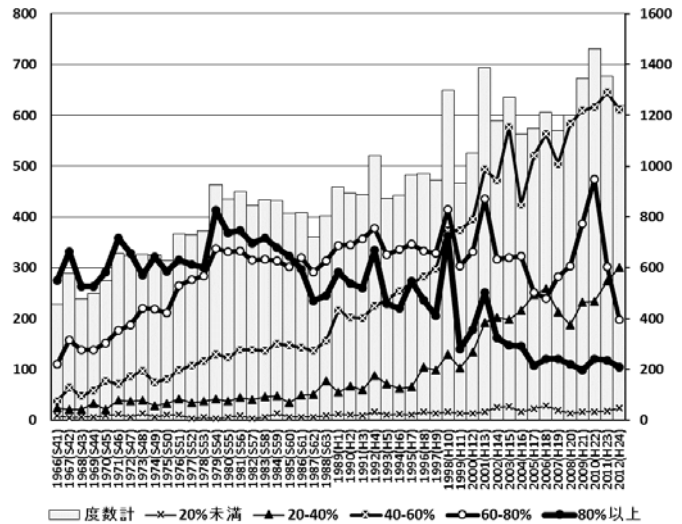
## 2. 各調査手法の回収率

### 2-1. 面接調査の回収率

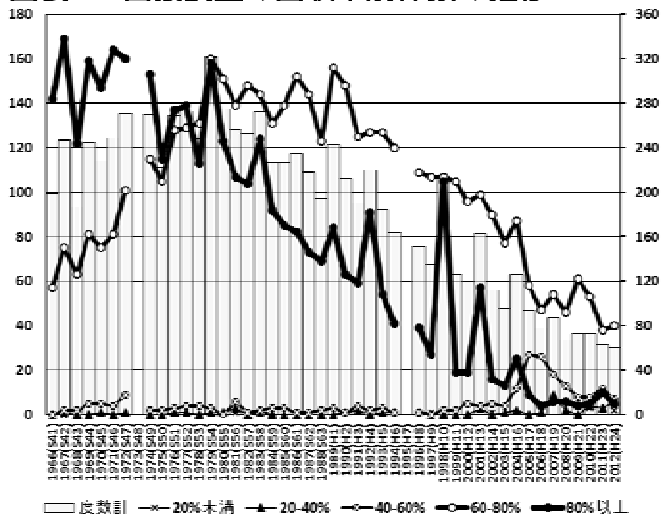
『世論調査年鑑』による面接調査の回収率推移(図表2)をみれば、回収率80%以上と60-80%の2曲線の減少傾向と比例して実施件数も減少している。

ただし、2005年近辺では、調査員の不正事件(日銀および内閣府が実施した調査で回収率70~75%から約50%に訂正された)の影響を受け、回収率60-80%曲線が急低下するのに合わせて50-60%曲線がピークを形成している。近年はこのピークも解消し、60-80%曲線の一時的な上昇や80%以上曲線の上向き傾向が読み取れる。公式統計をみれば、面接調査の回収率は継続的な低下傾向から脱しつつある。

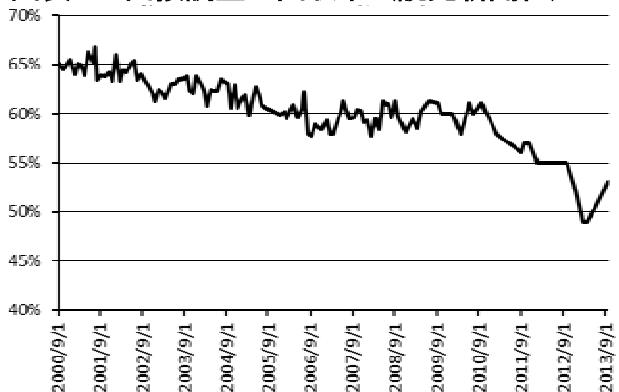
図表1 日本の世論調査の回収率別件数の推移



図表2 面接調査の回収率別件数の推移



図表3 面接調査の回収率(読売新聞社)



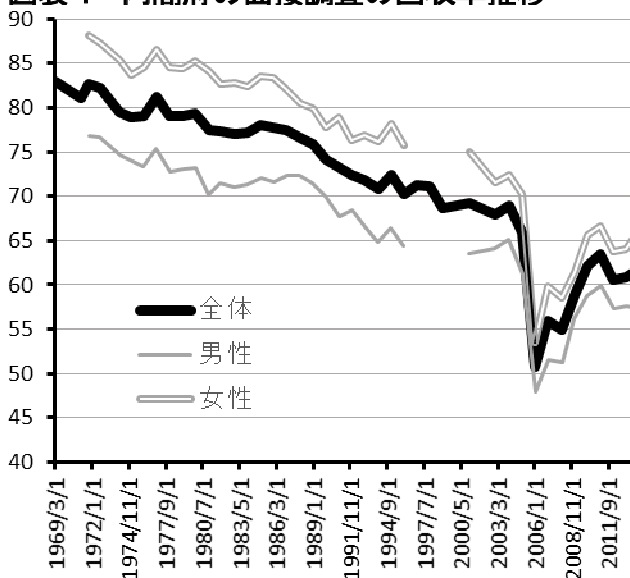
しかし、報道機関の実施する面接調査は正反対の傾向を示している。1978年から毎月（近年は年に4回程度）面接調査を実施してきた読売新聞社の回収率は、2010年から急減している。2013年2月、3月の両調査においては49%と半数を割っている。同期時（2013年3月）に実施された朝日新聞社の面接調査の回収率も51%に急低下（前回2011年12月は63%）している。この調査は、実は政府の調査を請け負っている調査会社に委託して実施された。前回までは、朝日新聞記者が学生を管理して実施していたが、調査会社に委託することで回収率は急落している。この結果を受けて、朝日新聞社は継続的に実施してきた面接調査を打ち切る決断をした。毎日新聞社も毎年の読書調査に合わせて実施していた面接調査を郵送調査に切り替えている。

一方、政府が実施する面接調査の回収率は、V字回復している。図表4は内閣府が実施する「社会意識に関する世論調査」の回収率推移である。2006年近辺での急落は先述した不正事件と個人情報保護法施行の影響を受けたと考えられる。しかし、報道機関の惨状とは違い、近年の回収率は60%超まで回復し、維持されている。

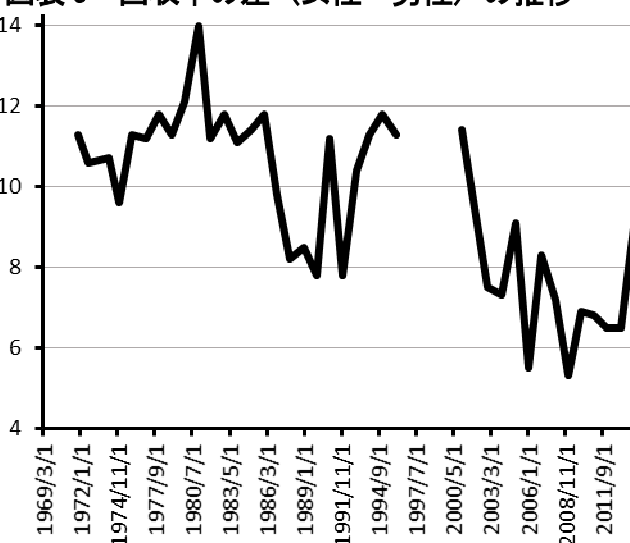
内閣府の調査仕様では、予備サンプルを使用していない。予備サンプルとは、抽出時点で存在していた対象者が調査時点で死亡や転居などにより調査不能となった場合の代替として用意しておくものである。予備サンプルの使用にあたっては、病気や長期不在に対しても適用するかどうか、さらに判断が甘くないのかどうかといった運用上の課題がある。先述した読売新聞社は予備サンプルを使用している（窪田，2008）。おそらく、内閣府のように予備サンプルを使用しないという基準で読売新聞社の回収率を計算すれば、さらに回収率が低くなり（40%台）、内閣府の面接調査と15～20%程度の差が確認されるであろう。

内閣府が実施する面接調査の回収率が向上していることは喜ばしいことである。半面、面接調査の回収率が報道機関と政府とで大きく乖離していることには大きな不安を感じざるを得ない。

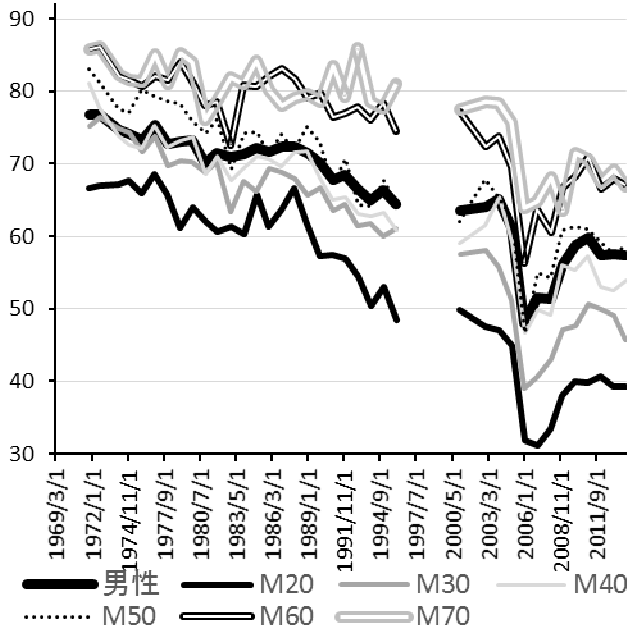
図表4 内閣府の面接調査の回収率推移



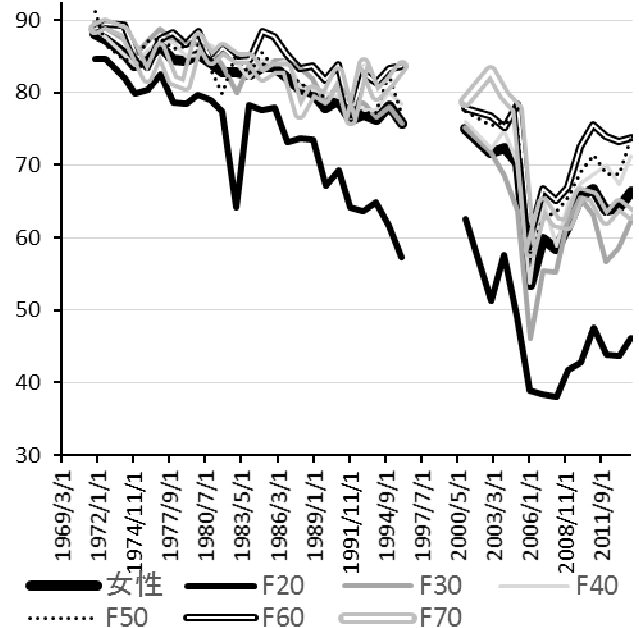
図表5 回収率の差（女性－男性）の推移



図表6 男性の年代別回収率の推移



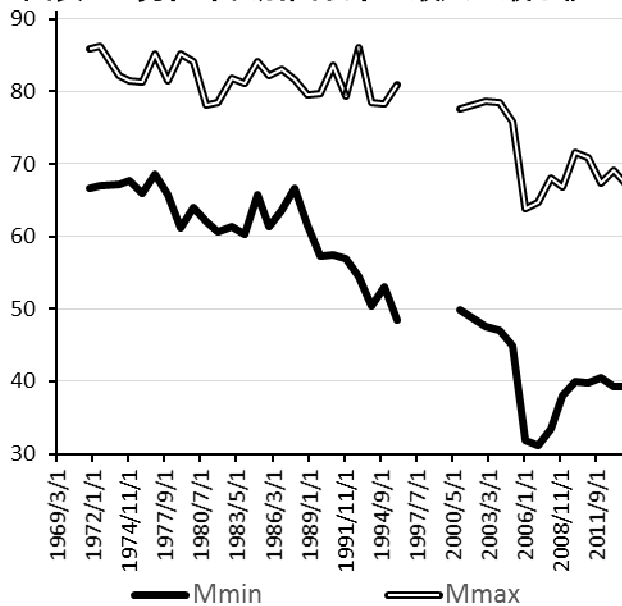
図表7 女性の年代別回収率の推移



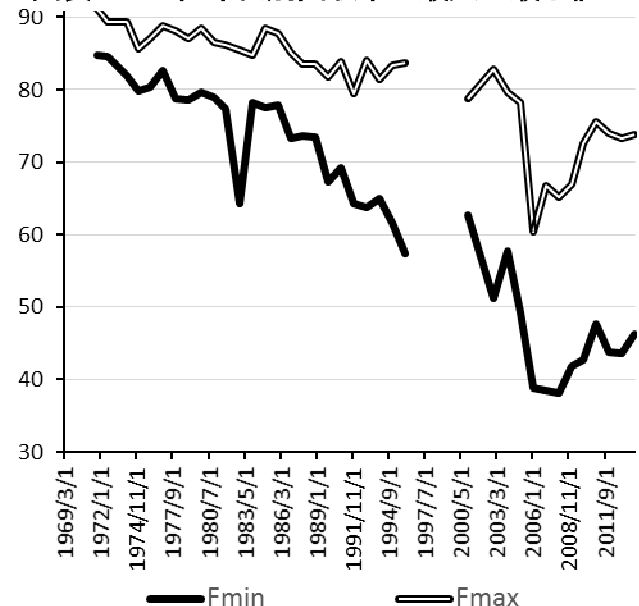
図表4では一貫して男性よりも女性での回収率が高いが、2006年近辺では女性の回収率が低下し男性のものに接近している(図表5)。近年の60%台回復に向けての回収率の変動は、男性よりは女性での回収率上昇の影響を受けているようである。

さらに男女別で、年代ごとの回収率の推移を見してみる(図表6、7)。男性では1970年代から近年まで、年代別回収率の最大値と最小値の差がやや大きくなってきている程度であるにもかかわらず、女性では差がかなり大きくなってきている。1971年では男性19.2%、女性6.4%と3倍の開きがあるが、2014年は男性27.9%、女性27.6%とほぼ同率である(図表8、9)。

図表8 男性年代別回収率の最大・最小値



図表9 女性年代別回収率の最大・最小値

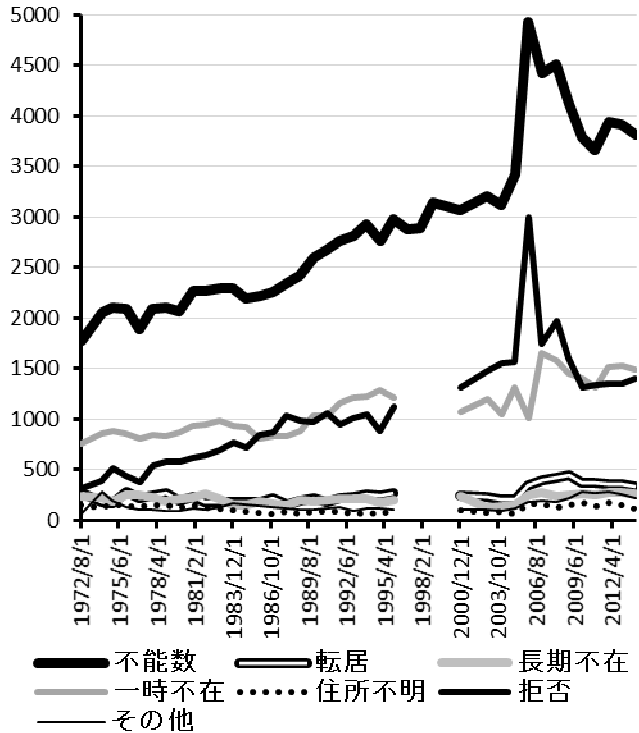


回収できなかった調査不能の数（不能数）と不能理由の内訳の推移を確認してみる（図表10、11）。不能数は回収数（率）の裏返しになるため、当然2006年近辺で鋭いピークになっている。不能理由の中で「拒否」が同じように鋭いピークをもつ。これは、当時の回収率低下（不能数増加）は個人情報保護法施行によるプライバシー意識の高まりの影響を受けたと指摘されていることを裏付ける現象といえる。

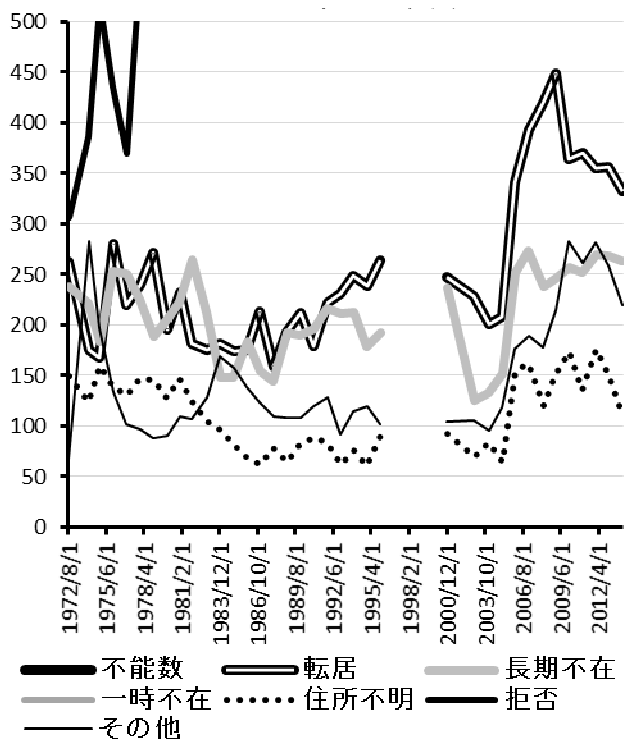
さらに、不能理由としては「拒否」や「一時不在」よりは度数が少ないが、「転居」と「その他」が2005年近辺から急増している。「転居」が急激に増えるということは、①抽出時期と調査時期に大きな開きがあったか、②名簿の更新が滞ったか、どちらかの可能性がある。個人情報保護法施行の影響を受けて、閲覧手続きに手間取り抽出期間が数カ月にわたるといった事態が発生したか、住民からの苦情対策として閲覧用抄録の更新を控えたか、当事者に確認しておく必要がある。

これらの不能理由の構成変化を見る限り、内閣府が実施する世論調査の回収率が急回復しているのは、名簿閲覧作業にかかわる回収率低下の誘因よりは、調査「拒否」が急減したことによる効果といえる。回収率の急落への対応として、①調査主体が内閣府だと明かして調査を依頼する、②回答者には謝礼を渡す、という運用改革を実施したことの成果といえる。しかし、統計数理研究所が実施した2013年「日本人の国民性」調査の回収率は49.5%と低い（参照：行動計量学会での発表）。政府調査の「拒否」が「一時不在」を下回っていることの誘因をさらに精査する必要があるのではないかと。

図表10 不能理由の推移（度数）



図表11 不能理由の推移（度数500以下拡大）



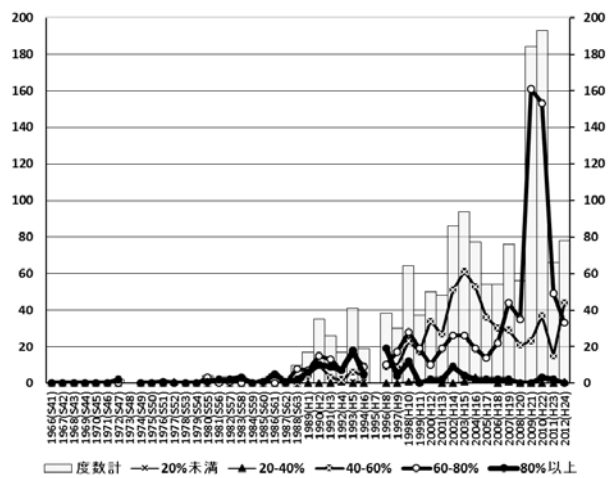
## 2-2. 電話調査の回収率

『世論調査年鑑』による電話調査の回収率推移(図表12)をみれば、回収率60-80%曲線が2009年(161件)と2010年(153件)で鋭いピークを持ち、面接調査の該当回収率曲線(図表2:2009年61件、2010年53件)を上回る。日本においては、面接調査よりも電話調査の回収率が高く、その該当件数も格段に多いという統計になっている。

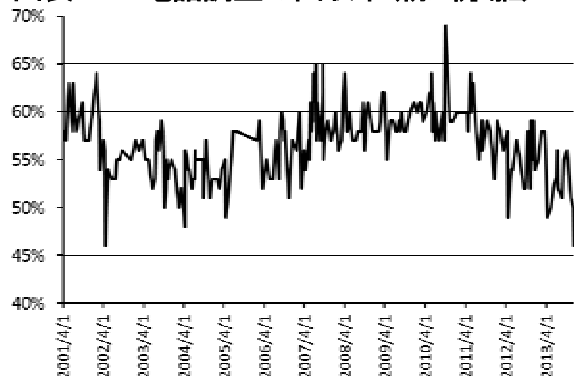
これは、将来に残る公式統計として問題がある。面接調査の回収率計算では計画標本数を分母にするのが基本であるが、予備サンプルを使用することで回収率が水増しされている事例が多い。それでも、電話調査の回収率水増しよりはまだましである。報道機関で実施されているRDD調査では、調査期間中に世帯用か事業用のどちらか判明しない電話番号を回収率計算の分母から取り除くという操作を行っている。調査終了後もこれらの番号に電話をかけ続けて世帯用か事業用かを判明させようと回収率を計算し直せば、実質回収率はかなり低くなる。過去の試行では、調査終了後に世帯か事業所か未判明だった番号に電話をかけ続けると、5割以上が世帯用と判明しているからである。報道機関はそのことを承知している。ただ、調査終了直後に報道する必要があるため、回収率計算の際に、分母を「世帯と判明した電話番号」として、その人たちに対する「回答率」として発表している。計画標本に対する「回収率」という表現を使用していない。つまり、『世論調査年鑑』は、面接調査などにおける「回収率」と電話調査などにおける「回答率」が混同した公式統計ということになる。

こうした「回答率」は、調査精度を反映する指標としては機能せず、むしろ運用実態を反映したものとなる。報道機関で一番多くRDD法による電話調査を実施してきた朝日新聞社の回収率推移で確認する(図表13:世論調査協会が開示しているデータより作成)。個人情報保護法施行後のプライバシー意識の高まりなどで調査の運用が難しくなっているにもかかわらず、2004年から2010年間のRDD調査の回収率は上昇傾向を示している。有効獲得サンプルの減少対策として計画標本(用意する電話番号)を増やし、未判明番号も増えるような運用をすれば目標有効数を確保すると同時に「回答率」(回収率ではない)も上昇する。一方で、2011年以降は回答率が減少

図表12 電話調査の回収率別件数



図表13 電話調査の回収率(朝日新聞社)



している。内閣府の面接調査の「回収率」がV字回復したのとは逆で、朝日新聞社の RDD 調査の「回答率」はΛ(ラムダ)字失墜している。これが、「回答率」は調査精度を反映していないが、運用に課題があることを明らかにしている事例である。

正確な「回収率」を確認しない運用を続けることで、RDD 調査が正常に機能しているのかどうか分からない現象が起きている。各社が報道する内閣支持率が大きく異なる事例のことである。これは「事例」ではなくもはや「事件」というべきものかもしれない。

2008年8月の福田内閣改造直後の内閣支持率は朝日・毎日と読売・日経が大きく異なっていた。このときは、質問文に「改造」という文言が加わっていることと回答を渋る人に「重ね聞き」している影響で、読売・日経の内閣支持率が高めにな

図表 14 報道各社の内閣支持率の比較(2008年と2014年)

	2008年	支持(前回比)	不支持	「改造」	重ね聞き	回答率(数)
読売	8/1-2金土	41.3(+14.7)	47	○	○	58(1006)
日経	8/2-3土日	38(+12)	49	○	○	61(856)
共同	8/1-2金土	31.5(+4.7)	48.1	○	×	71(1025)
毎日	8/1-2金土	25(+3)	52	×	×	60(921)
朝日	8/1-2金土	24(±0)	55	×	×	58(1002)
	2014年	支持(前回比)	不支持	「改造」	重ね聞き	回答率(数)
読売	9/3-4水木	64(+13)	29	×	○	55(1005)
日経	9/3-4水木	60(+11)	26	×	○	64(945)
共同	9/3-4水木	54.9(+5.1)	29.0	○	×	69(1016)
毎日	9/3-4水木	47(±0)	32	×	×	59(1037)
朝日	9/6-7土日	47(+5)	30	×	×	44(1614)

ったという識者のコメントが相次いだ。そして、今年(2014年)また、朝日・毎日と読売・日経で差が開いた。しかし、今回は読売・日経の質問文には「改造」の文言は入っていない。そのため、「重ね聞き」の影響だとする識者のコメントが増えている。しかし、松田(2010)は当時から、①第1次安倍内閣のときはこのような現象は発生していない、②2008年当時、読売新聞社の内閣支持率調査は面接調査が中心であり緊急時のみ RDD 調査を使用していた、③読売と日経新聞社の通常時の RDD 調査は3日間調査であり緊急時には1.5日調査となる、ことを根拠に一貫して運用の問題であることを指摘してきた。「重ね聞き」は内閣支持率が高めになる誘因にすぎず、その根本には緊急招集されるオペレーターの不均質な運用や改造内容を正確に把握していない調査1日目の夜に質問された対象者がどのような回答行動をとるのかという「どのように調べたものを内閣支持率とするかという概念」が起因となっているはずである。さらに、目標有効数を確保するため緊急時は通常時よりも電話番号サンプルを増やすが、調査期間の短さからコール回数は少なくなる。「回答率」の計算定義を思い起こせば、回答率は通常と同等かむしろ高くなる傾向がある。こうした、真の回収率定義をせずに回答率定義で運用される調査は、見た目の体裁が優先され、調査の質を落とし続けることになる。

### 2-3. 郵送調査の回収率

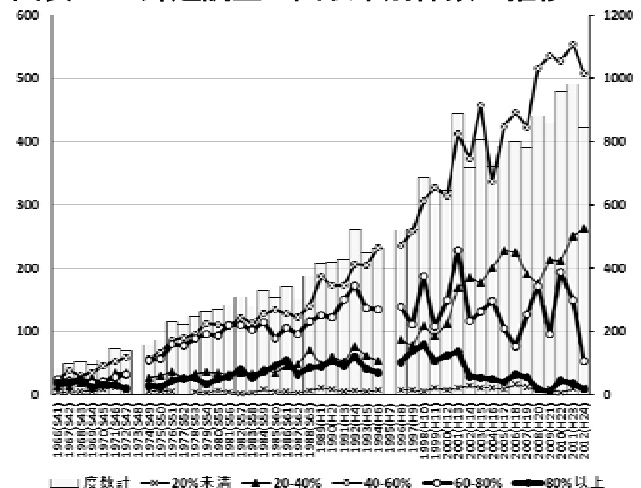
『世論調査年鑑』による郵送調査の回収率推移(図表15)をみれば、回収率40-60%曲線が1985年(134件)から60-80%曲線を突き放し2008年以降は500件を超えている。1996

年から 20-40% 曲線が傾きを変え急増していることが、郵送調査の実施件数の増加（傾き）にも反映されていることが分かる。

近年は 80% 曲線や 60-80% 曲線が急低下しているが、読売、朝日、毎日などの各新聞社が実施した郵送調査は 60% 以上を確保している。内閣府が 2014 年 3 月に「社会意識に関する世論調査」（面接法）の直後に並行的に実施した郵送調査の回収率も 75.3% である。「郵送調査の回収率は低い」という誤った常識に甘んじて回収率 40-60% の確保で満足したり、20-40% でも当然の結果だという認識による運用が、統計上に反映されているのではない。

日本においては、「回収率」の意義をはき違えていることが、健全な調査を損ない病状の早期発見を見逃す事態となっている。面接や電話調査における高回収率は、調査運用の異常を予感させ、低回収率が操作のない調査を予感させるという「回収率の逆指標性」がある。一方、回収率定義が明確なアメリカでは、別の展開が始まっている。

図表 15 郵送調査の回収率別件数の推移



### 3. アメリカの調査状況

新しい調査方法は、選挙予測の失敗と成功から生まれている。1936 年アメリカ大統領選挙でリテラー・ダイジェスト誌が予測に失敗したことで、「標本サイズの大きさよりは標本の代表性」を重視する流れに変わった。このとき、ギャラップ、ローパー、クロスレーは representative method（いまの割当法の一つ）を用いて成功したが、1948 年大統領選挙ではそろって失敗した。これにより、「代表性は割当ではなく無作為で担保する」という教義が出来上がった。私たちが社会調査法で教授される伝統的な無作為抽出法による調査は、この経験を受け継いでいる。さて、いまもこの教義は正しいのだろうか。

2000 年アメリカ大統領選挙でハリスインタラクティブがインターネット調査を用いて他社の電話 (RDD) 調査より 2 倍正確に予測的中させたときには、ギャラップなどの老舗調査会社や学術関係者は科学的ではない（理論に基づいていない）と批判した。固定電話を対象にした RDD 調査を比較基準とした傾向スコアを用いて、インターネット利用者でない有権者の意向も推定するいわゆるカバレッジ・エラーの補正をした。ハリスの方法は選挙予測の有力な方法論としては普及しなかったが、傾向スコアは、回収率が低下していた伝統的調査のノンレスポンス・エラーを補正する手段として多用されるようになった。

そして今年 7 月 25 日のニュースは、調査関係者に衝撃を与えた。NYT (New York Times) と CBS (CBS News) が共同で Battleground Tracker と銘打ち、YG (YouGov) が実施するインターネット調査を利用して 11 月まで毎月、中間選挙の予測報道を実施すると発表したの



である。第1回報道は7月27日。NYT、CBS、YGがともにネット上で共和党がやや強いという予測結果を掲載した。

伝統的調査を重視するAAPOR（アメリカ世論調査学会）の中心メンバーであり、重い報道責任を担うNYTとCBSの変節に対し、AAPORは8月1日付で「今日の選挙・世論調査における透明性と規格の重要性」と題した声明を公表している。

この声明に対する評判はよくない。Reg Baker（コンサルタント：AAPORやMR関連の役職を担い続けている）は「AAPORは踏み外した(AAPOR gets it wrong)」と題する記事を公表した。AAPOR声明の中にある「これら新生の方法論の限界を十分に理解することなく新しい試みとして取り込むことの性急さ」「透明性、完全開示、明確な規格という強力な枠組みの欠乏」という骨子に対し苦言を呈している。約30年AAPOR会員であり、特別委員会の共同議長も務めたと自己紹介しながら、AAPORは「戦略的な隠喩(political metaphor)を用いたことで歴史の悪役になってしまった」と嘆き、「未来を拒否するのではなく、世界中の有能なメソドロジストたちと一緒に、新しい方法へと移行する人達を援助することに力を注いでほしい」と主張している（Baker, 2014）。

Bakerは2012年アメリカ大統領選挙での各社の予測精度を示す一覧表（図表16:Silver, 2012）を引用している。今回問題視されたYG(YouGov)の平均誤差は2.6%なのに対し携帯電話も対象にしたGallupの電話調査は7.2%と大きく劣っている。YGのマッチング法は学会活動や論文でも開示されている（Rivers, 2006, 2007, & Bailey 2009）。

つまり、新しい方法ではなく伝統的な方法のほうに限界があること、このマッチング法は十分な透明性があり前途有望であることを示してAAPORの声明を批判したわけである。

Bakerの声明発表の前日にはGelman（コロンビア大学教授：統計・政治学の専門家でありベイズ統計の著書もある）とRothschild（エコノミスト：おもにMSR-NYCの仕事をしている／マイクロソフトリサーチ）が共同でThe Monkey CageとThe Huffington PostのコラムでAAPORの声明に対し反論している（Gelman/Rothschild, 2014）。

Baker同様にAAPOR声明の問題箇所を引用し、NYTのThe Upshotのコナーを担当しているCohn(2014)がYGを起用する明解な理由をすでに7月27日に説明していることを指摘し、

図表16 2012年アメリカ大統領選挙の予測精度

**Pollster Accuracy and Bias, 2012 Presidential Election**

Likely Voters Polls in Last 21 Days of Campaign  
Minimum 5 Polls

Pollster	# Polls	Avg. Error	Bias	Mode	Cell?
IBD / TIPP	11	0.9	R +0.1	Live Phone	Yes
Google Consumer Surveys	12	1.6	R +1.0	Internet	N/A
Mellman	9	1.6	R +0.0	Live Phone	Yes
RAND Corporation	17	1.8	D +1.5	Internet	N/A
CNN / Opinion Research	10	1.9	R +0.6	Live Phone	Yes
Ipsos / Reuters (online)	42	1.9	R +1.4	Internet	N/A
Angus Reid	11	1.9	R +0.8	Internet	N/A
CVOTER International / UPI	13	2.0	R +2.0	Live Phone	??
Grove Insight	18	2.0	R +0.1	Live Phone	Yes
SurveyUSA	17	2.2	R +0.5	Robodial	Yes
Quinnipiac	5	2.3	D +0.3	Live Phone	Yes
Marist	11	2.5	R +1.0	Live Phone	Yes
YouGov	30	2.6	R +1.1	Internet	N/A
We Ask America	9	2.6	D +0.1	Robodial	No
Public Policy Polling	71	2.7	R +1.6	Robodial	No
Gravis Marketing	16	2.7	R +2.7	Robodial	No
JZ Analytics*	17	2.8	R +0.1	Internet	N/A
Washington Post / ABC News	16	2.8	R +2.7	Live Phone	Yes
Pharos Research Group*	14	4.0	D +2.5	Live Phone	No
Rasmussen Reports	60	4.2	R +3.7	Robo + Internet	No
American Research Group	9	4.5	R +4.5	Live Phone	Yes
Mason-Dixon	8	5.4	R +2.2	Live Phone	Yes
Gallup	11	7.2	R +7.2	Live Phone	Yes

\* Not used in FiveThirtyEight forecast.

AAPORの妨害行為を揶揄している。

Gelmanの反論は明解である。伝統的調査が用いている仮説(①電話で誰にでもつながる、②どんな人も電話では同じように回答する、③どんな人でも同じように調査に参加する、④調査者は一般有権者の回答から投票者を推定できる)が、妥当なのかどうか。携帯電話の普及による影響とPew Research Centerが開示した回収率の低下(図表17: Pew, 2012)を根拠に、伝統的調査が用いる仮説はYGのネット調査が用いる仮説と同じような状況になっていると強調している。

今回の中間選挙報道の議論から、「回収率の低い調査に対する科学性をどのように確保するのか」と

いう課題に対する視点の違いが明確になってきた。ここでの立場は、以下の2つである。

- ①**事後対応** 確率論を用いてサンプルの代表性を確保しても回収率が悪いので、有効回答を利用して後から補正する(課題→補正がきちんとできるのか?アメリカのRDDの回収率は10%以下!?)
- ②**事前対応** 回収率が悪ければ事後補正しても正確に直せないなので、本来の調査対象者に近い人を事前に属性や性向からマッチングして、本来の対象者に近い意見として利用する(課題→マッチングの変数が回答の代表性を保証するのか?)

スカートの丈や色の流行のように、歴史は螺旋状に技術の進歩を取り込みつつ概念の遷移を繰り返す。YGが用いる手法は、1936年にギャラップたちが開発したRepresentative Methodの現代版である。その線上で進化するのか、別の確率論に従う新しい伝統的調査が生み出されるのか。明確に回収率の低下を認識できるアメリカの調査界では、興味深い調査方法論の議論が展開している。

#### 4. 終わりに

回収率は、調査の偏りを見極める指標である。しかし、明確な定義が周知されていない日本では、偏りを反映しない指標となり、調査手法の劣化が見えにくくなっている。調査の危機的状況の解決は、その病状を明確にすることから始まる。まずは厳格な回収率定義とその周知が求められる。(埼玉大学社会調査研究センター准教授)

#### 謝辞

Riversの論文はMokrzycki, Mike (NBC News 選挙調査部長)がツイッターで周知したものである。感謝する。

図表17 Pew Research Center 開示の回収率推移

	1997	2000	2003	2006	2009	2012
	%	%	%	%	%	%
<b>Contact rate</b> (percent of households in which an adult was reached)	90	77	79	73	72	62
<b>Cooperation rate</b> (percent of households contacted that yielded an interview)	43	40	34	31	21	14
<b>Response rate</b> (percent of households sampled that yielded an interview)	36	28	25	21	15	9

PEW RESEARCH CENTER 2012 Methodology Study. Rates computed according to American Association for Public Opinion Research (AAPOR) standard definitions for CON2, COOP3 and RR3. Rates are typical for surveys conducted in each year.

〈参考文献〉

- 松田映二「電話調査の環境変化と手法遷移」『行動計量学会第38回大会抄録集』、162-163、2010
- 松田映二「調査の信頼性を取り戻すために」『政策と調査』6、3-37、2014
- 窪田和久「面接調査の現状と課題」『行動計量学』35(1)、5-16、2008
- Baker, Reg (2014). AAPOR gets it wrong:  
[http://regbaker.typepad.com/regs\\_blog/2014/08/aapor-gets-it-wrong.html](http://regbaker.typepad.com/regs_blog/2014/08/aapor-gets-it-wrong.html) (August 05, 2014)
- Cohn, Nate (2014). Explaining Online Panels and the 2014 Midterms  
<http://www.nytimes.com/2014/07/28/upshot/explaining-online-panels-and-the-2014-midterms.html?module=Search&mabReward=relbias%3Aw%2C%7B%22%22%3A%22RI%3A8%22%7D&abt=0002&abg=1> (JULY 27, 2014)
- Gelman, Andrew and David Rothschild (2014). Modern polling needs innovation, not traditionalism  
<http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/08/04/modern-polling-requires-both-sampling-and-adjustment/> (August 4, 2014)
- Pew Research Center (2012). Assessing the Representativeness of Public Opinion Surveys  
<http://www.people-press.org/2012/05/15/assessing-the-representativeness-of-public-opinion-surveys/> (May 15, 2012)
- Rivers, Douglas (2006). Sample Matching: Representative Sampling from Internet Panels. Palo Alto, CA: Polimetrix, Inc.  
[http://web.mit.edu/clawson/www/polisci/OldFiles/portl/cces/material/sample\\_matching.pdf](http://web.mit.edu/clawson/www/polisci/OldFiles/portl/cces/material/sample_matching.pdf)
- Rivers, Douglas (2007). Sampling for Web Surveys. Paper presented at the 2007 Joint Statistical Meetings, Salt Lake City, Utah, July-August.  
[http://www.laits.utexas.edu/txp\\_media/html/poll/files/Rivers\\_matching.pdf](http://www.laits.utexas.edu/txp_media/html/poll/files/Rivers_matching.pdf)
- Rivers, Douglas, and Delia Bailey (2009). Inferences from Matched Samples in the 2008 U.S. National Elections. Proceedings of the Survey Research Methods Section of the American Statistical Association, Joint Statistical Meeting 2009.  
<https://www.amstat.org/sections/srms/proceedings/y2009/Files/303309.pdf>
- Rothschild, David and Andrew Gelman (2014). Modern Polling Requires Both Sampling and Adjustment  
[http://www.huffingtonpost.com/david-rothschild/modern-polling-requires-b\\_b\\_5646174.html](http://www.huffingtonpost.com/david-rothschild/modern-polling-requires-b_b_5646174.html)  
(Posted: 08/04/2014 12:40 pm Updated: 3:59 pm)
- Silver, Nate (2012). Which Polls Fared Best (and Worst) in the 2012 Presidential Race  
[http://fivethirtyeight.blogs.nytimes.com/2012/11/10/which-polls-fared-best-and-worst-in-the-2012-presidential-race/?\\_php=true&\\_type=blogs&\\_r=0](http://fivethirtyeight.blogs.nytimes.com/2012/11/10/which-polls-fared-best-and-worst-in-the-2012-presidential-race/?_php=true&_type=blogs&_r=0) (November 10, 2012 8:38 pm)