

米国における RDD サンプルングの最近の動向

島田喜郎
株式会社日経リサーチ

1 まえがき

米国では Waksberg 法は RDD サンプルングの「標準」の手法と言う地位を与えられていた。しかし最近、Waksberg 法はあまり使われていないという印象がある。「Waksberg 法はどうなったのか」という関心を中心にして、米国の最近の RDD サンプルング手法の動向を、理論的な展開と現実的に調査に使われる手法の両面から概観する。

2 理論的背景

RDD サンプルングでは、固定電話加入世帯を母集団として、電話番号を世帯 ID に使ってサンプルングを行う。世帯用電話番号の完全なリストが無いので、世帯用電話番号を完全に含む（あるいはほぼ完全に含む）抽出フレームを定義する。この抽出フレームから無作為抽出して RDD 標本を作り、RDD 標本の中の世帯用番号を電話加入世帯の標本として使う。抽出フレームには稼働局番フレームと list-assisted フレームがある。

稼働局番フレームは現に稼働中のすべての局番の 0000-9999 の番号が抽出の対象である。このフレームは世帯用番号を完全に包含している。稼働局番フレームから無作為抽出した RDD 標本の世帯ヒット率は日本でも、米国でも 20%を少し上回る程度で、このヒット率の低さが問題とされる。

list-assisted フレームは、電話帳(list)に掲載されている世帯用番号を少なくとも 1 件以上、あるいは一定件数以上もつバンク(bank)を集めたものである。バンクは局番の中を一定の大きさのブロックに分けたもので、012-345-67xx のように局番

の先頭から百位までの数字が同じ 100 個の番号のブロックを百位バンクと呼ぶ。千位バンクや十位バンクも考えられるが、よく使われるのは百位バンクである。百位バンクを使った RDD 標本の世帯ヒット率は 50%を上回る。list-assisted フレームは世帯用番号がすべて非掲載であるバンクを排除してしまうので、別名脚きり(truncated)フレームとも言われ、電話世帯のカバレッジが完全でないことが問題とされる。

Waksberg 1978 の方法は電話世帯を完全にカバーしながら list-assisted 法に劣らぬヒット率を達成した。稼働局番フレームから一次抽出単位(PSU)として百位バンクを抽出し、2 段目にその百位バンクから一定数の世帯用番号を抽出する 2 段抽出法である(紙面の都合で Waksberg 法をこれ以上詳しく説明できないがご容赦願いたい)。この手法はそれまで電話調査を避けていた政府機関などの要求水準の高い調査にも採用されるようになった。Groves et al(編) 1988 は当時の電話調査の理論的な集大成ともいえる文献だが、Waksberg 法はその中で RDD サンプルング手法の主役の立場を与えられた。

しかし、Waksberg 法にも問題があった。最大の問題は、サンプルングとフィールドワークの作業が絡み合うため、手順が複雑で時間を要することであった。1990 年代になって発表された論文が Waksberg 法の命運に大きな影響を与えることになった。

Casady&Lepkowski 1993 は、list-assisted フレームを使いながら、掲載番号の無いバンクからも抽出する方法を提案した。1 以上の掲載番号をもつ 1+バンク層と掲載番号の無い 0 バンク層に

わけ、層によって抽出率を変える。たとえば、0バンク層での抽出率を1+バンク層の3分の1とか4分の1にする。非比例層化抽出法 (disproportionate stratified sampling)である。この手法は電話世帯を完全にカバーしながら、段階の抽出で世帯番号に到達する単純さと高い世帯ヒット率を同時に達成することを可能にした。

一方、Brick et al 1995 は、list-assisted 法で脚きりされるバンクに属する世帯が電話世帯の3-4%程度に過ぎず、その世帯は抽出フレームに含まれる世帯と比べてさまざまな特性において有意な差が見られないことを確認した。その結果、脚きりによって生じるバイアスは特に重視するほどのものでないと断定した。Giesbrecht 1996 もほぼ同様の調査結果を発表した。

3 米国のRDD調査の現状

実際に米国では現在どんな手法で RDD サンプルングが行われているのか、インターネット上のドキュメントを利用して検証した。検索サイト Google で「RDD sample survey methodology Waksberg」をキーワードとして検索した。その結果から「実際に行われた調査の報告で、RDD サンプルングの手法が明記されているもの」だけを抜き出したのが表1である。検索語に Waksberg を含めたのは、できるだけ Waksberg 法が拾えるように考えたためである。

表1によると、Waksberg 法を使っているものは25件中5件に過ぎなかった。Waksberg 法以外の方法は、「単純 RDD」が1件で、他はすべて list-assisted 法であった。Waksberg 法を使わなかった20件のうち5件では、Waksberg 法を「過去に使用」したという記述があった。また8件では「標準」の Waksberg 法を使わない理由として Waksberg 法の「問題点を指摘」したり、新しい方法の相対的な利点をあげていた。

list-assisted 法で抽出フレームに含めるバンクの条件は、掲載件数を1以上としたものが多かったが、2以上、3以上などのケースも見られた。

RDD 標本を提供した会社に言及しているケースが多く、サンプリング会社の存在が list-assisted 法の普及に大きな役割を果たしていると言えそうである。

表1で Waksberg 法を使った5件の調査がすべて「健康」関連の調査であることが注目される。これには BRFSS (Behavioral Risk Factor Surveillance System) の影響が考えられる。BRFSS は、米国保健省の指導の下に各州が主体になって実施する生活習慣病の調査である。BRFSS は標本設計のルールとして電話世帯を完全にカバーすることを要求している。表1の Waksberg 法の使用例は、BRFSS 調査であるものと BRFSS から何らかの影響を受けているものが含まれていると思われる。しかし、BRFSS が Waksberg

法で固まっていると考えるのは間違いない。表

表2 BRFSS の RDD 手法採用州数

	1995年	2000年
非比例層化抽出法	12	50
Waksberg 法	34	1
単純 RDD 法	1	1
その他	4	1
計	50	53

1995年は Iachan, R. et al. 2001 による
2000年は BRFSS のウェブサイトによる

2によると、BRFSS のサンプリング手法は1995年に主流であった Waksberg 法が、2000年には完全に後退し、ほとんどの州が非比例層化抽出法を使うようになった。

4 まとめと考察

米国において、一時期、RDD サンプルングの標準の手法と考えられていた Waksberg 法は実質的に消え去りつつあるといえる。Waksberg 法に代わって頻繁に使われるようになったのは、list-assisted フレームを使う方法である。list-assisted 法による「脚切り」を容認せず電話世帯を完全にカバーすることにこだわる場合には、非比例層化抽出法や単純 RDD 法が使われている。

list-assisted 法を採用したケースでは、Brick et al. 1995 や Giesbrecht et al. 1996 などの文献に言及するものが多く、これらの文献がこの手法を

正当化するために重要な役割を果たしているといえる。両文献とも、掲載件数1以上の百位バンクのフレームで検証を行ったにもかかわらず、掲載件数の条件が2以上や3以上のケースでもこれらを根拠文献として「便利」に引用しているのは意外である。今後ともこれらの文献は便利に引用されると思われる。電話帳掲載状況は時代とともに変化するものであり、Brick 等が確認したのはその一時点の状況である。いつまでもそれを根拠にすることが妥当であるとは思えない。

非比例層化抽出法は電話帳情報を利用しながら、電話世帯を完全にカバーするので、電話帳掲載状況の変化に耐えやすいし、米国とは異なる状況にも適用しやすい方法である。0バンク層の標本は抽出率を低くするため抽出件数が相対的に低く抑えられるとはいえ、この層のヒット率は非常に低く(たとえば、Casady & Lepkowski 1993 の場合 2%)、実査段階の効率を大幅に低下させると想像しがちである。しかし、標本抽出直後に自動スクリーニングを行って、使われていない番号や一部の事業所用番号の処理分類(disposition)を確定してしまえば効率にそれほど大きな影響を与えないはずだ。

自動スクリーニングは現在、米国でも日本でも RDD 調査で必ず使われているのに、スクリーニングとサンプリング手法の関係が論じられることが無いのは不思議である。稼動局番フレームを使う単純 RDD 法は、世帯ヒット率が低く非効率の見本のように考えられてきた。しかし、自動スクリーニングを使えば、単純 RDD 法でも効率のよい調査を実施することができる。表1と表2であげた米国の調査の中にも単純 RDD を使ったものがあった。それらが自動スクリーニング使ったどうかは明らかでないが、使ったと考えたほうが自然だろう。日本においては、日経リサーチが、2002年から単純 RDD 法と自動スクリーニングを組み合わせて、十分に効率的な RDD 調査を行っている。日経リサーチでは、稼動局番フレームから系統抽出した標本に対して、「使われていない番号」だけのスクリーニングを行う。非使用番号の比率

は RDD 標本の 63%にものぼり、世帯ヒット率は処理前の 2.7 倍に高まる。Casady & Lepkowski 1993 は、非世帯用番号であることを確定するコストと世帯のインタビューを完了するコストの比を使って層別抽出率の最適化を考えたが、電話帳情報を使うためのデータベースの作成・維持費用を考慮していない。また自動スクリーニングによって非使用番号の処理が安価にできることも考慮していない。それらを考慮した場合、1+バンクと0バンクで等しい抽出率を適用しても損はないという結論になる可能性があると思う。そうなれば、電話帳情報を利用する意味はほとんどなくなってしまふ。これは興味深い研究課題であると思う。

5 参考文献

- Brick, J. et al. (1995), "Bias in List-Assisted Telephone Samples", *Public Opinion Quarterly*, Vol. 59, No. 2, pp218-235.
- Casady, R. and Lepkowski, J. (1993), "Stratified Telephone Survey Designs", *Survey Methodology*, Vol 19, No 1, pp103-113.
- Giesbrecht, L. et al. (1996), "Estimating Coverage Bias in RDD Samples with Current Population Survey Data", in *Proceedings of the Survey Research Method Section*, American Statistical Association.
- Groves, R. et al. (Eds.) (1988), *Telephone Survey Methodology*, John Wiley & Sons.
- Iachan, R. et al. (2001), "Pooling State Telephone Survey Health Data for National Estimates: The CDC Behavioral Risk Factor Surveillance System, 1995" in Cynamon, L. et al. (editors), *Seventh Conference on Health Survey Research Method*, Department of Health and Human Services Centers for Disease Control and Prevention.
- Waksberg, J. (1978), "Sampling Methods for Random Digit Dialing," *Journal of the American Statistical Association*, Vol. 73, No. 367, pp. 40-46.

表1 インターネットで公開されている米国内調査の報告書における RDD サンプル抽出法

整理番号	調査名	実施年	主体名	抽出法	バンク基準	会社 Sampling	Waksberg 法		参照文献
							過去使用	問題指摘	
1	Prevalence of Diabetes among African Americans	1996	Missouri Department of Health	Waksberg					
2	Modifiable Cardiovascular Disease Risk Factors	1996	Missouri Department of Health	Waksberg					
3	Household Telephone Survey of the General Adult Population	1996	Oklahoma Department of Mental Health	Waksberg					
4	West Virginia BRFSS	1997	West Virginia Bureau for Public Health	Waksberg					
5	Massachusetts Health Benchmarks Polls	2002	University of Massachusetts Poll	Waksberg					
6	California Behavioral Risk Factor Survey (BRFS)	1995	California Department of Health Services	simple RDD			Yes		
7	National Election Surveys (non-time series study)	1992	Institute for Social Research, University of Michigan	list-assisted	2+	GENESYS	Yes		
8	Youth Attitude Tracking Study	1996	Department of Defense	list-assisted			Yes	Yes	
9	National Household Education Survey	1996	U.S. Department of Education	list-assisted	1+	GENESYS	Yes	Yes	CB
10	National Survey of Speeding and Other Unsafe Driving Actions	1997	NHTSA, U.S. Department of Transportation	list-assisted					B
11	Survey of Public Participation in the Arts	1997	National Endowments for Arts	list-assisted	1+				B
12	Massachusetts Adult Tobacco Survey	1997	Massachusetts Department of Public Health	list-assisted	1+	GENESYS	Yes		
13	Motor Vehicle Occupant Safety Survey	1998	NHTSA, U.S. Department of Transportation	list-assisted					B
14	Gambling Impact and Behavior Study	1998	National Opinion Research Center, University of Chicago	list-assisted	1+	SSI		Yes	B
15	National Survey of America's Families	1999	Urban Institute	list-assisted	1+			Yes	CBG
16	Y2K and the Banking Industry	1999	Board of Governors of the Federal Reserve System など	list-assisted	3+	SSI			CBG
17	National Immunization Survey	1999	Centers for Disease Control and Prevention	list-assisted	1+				
18	Secondhand Effect of Student Alcohol Use	1999	?	list-assisted		GENESYS			B
19	Driver Experiences and Expectations Regarding Conventional Brakes versus ABS	1999	U.S. Department of Transportation	list-assisted		GENESYS		Yes	B
20	BTS Omnibus Survey	2000	BTS, U.S. Department of Transportation	list-assisted	1+			Yes	B
21	Survey of Insurance Status	2000	Massachusetts Division of Health Care Finance and Policy	list-assisted		GENESYS		Yes	
22	Household Survey on Deposit Insurance Awareness	2000	Federal Deposit Insurance Corporation	list-assisted	3+	SSI			CBG
23	California Health Interview Survey	2001	California Department of Health Services など	list-assisted	1+	GENESYS		Yes	CBG
24	OHCA Household Survey	2001	Connecticut Office of Health Care Access	list-assisted	2+	GENESYS			B
25	Surveys of Consumers	2003	Survey Research Center, University of Michigan	list-assisted	1+	GENESYS			

検索語: rdd sample survey methodology Waksberg

検索エンジン: Google

掲載基準: 1) 実施された調査に関する報告で、2) RDD のどの手法を使ったか具体的な記述があるもののみを取り上げた。同じ調査が時系列で実施されている場合はその最新年のものを取り上げた

省略形

参照文献

BRFSS: Behavioral Risk Factor Surveillance System

BTS: Bureau of Transportation Statistics

NHTSA: National Highway Traffic Safety Administration

C: Casady and Lepkowski, 1993

B: Brick et al. 1995

G: Giesbrecht et al. 1996

発表要旨

著者： 島田喜郎
所属： 株式会社日経リサーチ
発表演題名： 米国における RDD サンプルングの最近の動向
発表要旨：

米国で RDD の標準手法ともいわれた Waksberg 法が、最近あまり使われていないという印象を検証するため、米国の最近の RDD サンプルング手法の動向を、理論的な展開と現実に調査に使われる手法の両面から概観した。たしかに Waksberg 法は使われなくなった。list-assisted フレーム法が盛んに使われている。この方法は「脚きり」によるカバレッジの損失があるが、それによるバイアスは重大では無いと言われる。しかし、電話世帯を完全にカバーすることを重視する人々は、掲載件数0のバンクからも抽出する手法や単純RDDを使うようである。

以上