

「調査への指向性」変数を用いた調査不能バイアス補正の試み

—「日本人の国民性調査」データへの適用—

土屋 隆 裕*

A Study of Nonresponse Adjustment with Information on Respondents' Survey Oriented Characteristics

—An Application to the Survey of the Japanese National Character—

Takahiro Tsuchiya

日本人の国民性第11次全国調査とその事後調査を用いて、調査不能バイアスの適切なウェイト補正の方法を検討する。まず、補正のためのウェイトはキャリブレーションにより求めるのがよいこと、ただし補助変数として人口統計学的属性変数のみを使うと、かえってバイアスが拡大する可能性があることを示す。次に、調査協力の得やすさ、あるいは回答者の調査への指向性を補助変数として加えることでバイアスをより低減できることを示し、指向性変数が持つべき特性を探る。最後に、第11次全国調査結果の補正を試み、補正の必要性の有無を検討する。

This paper explores four weighting nonresponse adjustment methods to determine their effectiveness. For this investigation, the postsurvey of the 11th nationwide survey of the Japanese National Character is used. First, it is illustrated that the calibration approach is preferable to the other three methods. However, it is possible that the calibrated weights may increase the bias when the auxiliary variables are limited only to demographic characteristics. Second, the paper demonstrates the effect of additional information from respondents' answers to attitudinal survey questions. Third, based on this investigation it is recommended that the results of this nationwide survey should be weighted.

Key Words and Phrases: nonresponse in social survey, calibration estimator, propensity score.

1. はじめに

統計数理研究所が1953年以来5年ごとに実施している「日本人の国民性調査」の回収率は、1980年代前半までは70%以上を保っていたが、その後急落し、2003年の第11次調査（坂元・中村・前田・土屋，2004）では56%にまで落ち込んだ。表1のとおり、回収率低下の主因は、拒否と一時不在の増加にある。

計画標本が自動加重である国民性調査では、回収標本における、各回答選択肢の単純な比率（標本比率）のみを公表している。しかし2003年には不能理由の半数近くを拒否が占め、母集団比率の推定量として、標本比率は大きなバイアスを抱えている可能性がある。例えば「#2.12c¹⁾人は信頼できるか」では、「信頼できると思う」という回答が1978年の26%から

* 情報・システム研究機構 統計数理研究所 データ科学研究系, 〒106-8569 東京都港区南麻布4-6-7

表1 「日本人の国民性調査」の回収率と調査不能の内訳

調査年	回収率 (%)	調査不能の内訳 (%)								計 (実数)
		死亡	移転	尋ね当 たらず	病気・ 老衰	長期 不在	一時 不在	拒否	その他	
1953年	83	1	3	1	4	2	5	1	1	101 (2,254)
1958年	79	0	6	2	3	3	4	2	1	100 (2,991)
1963年	75	0	7	3	3	5	3	3	1	100 (3,600)
1968年	76	0	7	2	3	4	5	3	0	100 (4,000)
1973年	76	0	6	2	2	5	4	4	0	99 (6,055)
1978年	73	0	5	2	3	5	6	6	1	101 (5,400)
1983年	74	0	2	2	3	2	8	9	1	101 (6,000)
1988年	61	0	3	2	5	3	11	15	0	101 (6,000)
1993年	69	0	3	2	2	2	10	13	0	101 (5,400)
1998年	64	0	4	1	2	1	9	17	1	99 (4,200)
2003年	56	0	4	2	3	2	12	21	0	100 (4,193)

2003年の33%へと7ポイント増加し、「用心するにこしたことはないと思う」は68%から59%へ9ポイント減少した。また「#2.12 他人のためか自分のためか」では、「他人の役にたとうとしている」が19% (1978年) から34% (2003年) へ15ポイント上昇し、「自分のことだけに気をくばっている」は74%から51%へ23ポイントも下落している。他人への信頼感が近年ほど増しているという、一見首肯し難いこれらの結果は、「用心するにこしたことはないと思う」人々が調査を拒否し、善意の人々の考え方だけが取り出されてきたためではないか、とも考えられる。

そこで本稿では、調査不能バイアスへの対処として、ウェイトを用いた補正 (Bethlehem, 2002; Gelman and Carlin, 2002) の可能性を探る。具体的には、以下の三つの目的に沿った分析を行い、補正の是非を実証的に検討する。第一は、第11次調査とその事後調査データを利用し、ウェイトを求めるいくつかの方法間で、補正効果の大きさを比較することである。Särndal and Swensson (1987) は、シミュレーションの結果、二相抽出法における二相目の調査不能バイアス低減には、回収率の逆数で単純に拡大するよりも、補助変数を用いた回帰推定量の方がよいと述べている。こういった知見が現実に通用的なのか、また実際に補正する際の留意点については、現実の調査では“真の値”を知りようがなく、傍証を積み重ねる必要がある。事後調査は第11次調査の回収標本に対し実施した郵送調査である。その返送者データを基に回収標本全体を推定し、採用すべき方法について指針を得る。土屋 (2005) は、この事後調査データを用いて返送者から未返送者の推定を試みており、本稿はその再分析と位置づけられる。なお、不能バイアスの検討にあたっては、第11次調査の計画標本全体に対して事後調査を実施する方が、多くの情報を得られることは論を俟たない。対象を回収標本に限定したのは、不能標本の調査に対する忌避感の助長回避やコストといった実査上の制約による。

第二は、「調査への指向性」変数を用いる効果を検証することである。一般に、性や年齢などの人口統計学的属性だけでウェイトを求めることは好ましくない (杉山, 1984) と考えられている。調査協力の得やすさ、あるいは回答者の調査に対する指向性の強さを補助変数に用いることの是非を、事後調査データを用いて検討する。

¹⁾ 「日本人の国民性調査」では#番号によって、調査項目を整理している。

第三は、以上で得られる知見を基に第 11 次調査結果の補正を試み、効果の大きさを検証することである。Ekholm and Laaksonen (1991) は、フィンランドの家計調査データを基にした世帯数の推定値が、補正の結果、標準誤差の 3.5 倍変わること示している。一方、国民性調査のような社会調査・意識調査では、補正による変化が数ポイント程度では、結果の解釈や結論に影響しないことが多い。解釈が変わり得るほどの大きな違いが見られるのか、実際の調査データを基に検証する。

本稿の構成は以下のとおりである。まず第 2 章では、第 11 次調査とその事後調査の概要を紹介する。第 3 章では、本稿で検討するウェイト集計の方法を述べる。第 4 章では事後調査データを用いた検証を行い、第 5 章で第 11 次調査結果の補正を試みる。最後に第 6 章でウェイト補正の可能性について考察する。

2. 日本人の国民性第 11 次全国調査とその事後調査の概要

2.1 日本人の国民性第 11 次全国調査の概要

調査対象：全国の 20 歳以上 80 歳未満の有権者。標本の抽出は層化二段抽出法による。層は区部・人口 20 万人以上の市部・20 万人未満の市部・郡部・沖縄県の 5 つ²⁾であり、比例割当とした。第一段として合計 300 投票区（地点）を確率比例抽出し（抽出率は層によって 0.2~1.4%）、第二段として地点ごとに平均 14 人の個人を選挙人名簿から等間隔抽出した。計画標本は全体で 4,193 名。

調査方法：個別面接聴取法。実査は社団法人 中央調査社に委託した。

調査内容：性・年齢などの人口統計学的属性の他、日本人のものの考え方を探るような質問項目。‘#2.7 一番大切なもの’を除く全ての項目がプリコード項目であり、#2.7 も事後にコード化した。用意した回答選択肢以外の答えは「その他」、質問の意味が分からなかったり答えられなかった場合は「D.K. (Don't Know; わからない)」とした。調査票は K 型と M 型の 2 種類あり、地点ごとに、名簿順に標本を K 型と M 型へ交互に割り当てた。

調査時期：2003 年 9 月下旬から 10 月上旬。

回収状況：表 2 のとおり、回収標本は全体で 2,350 名（回収率 56.0%）であった。

2.2 事後調査の概要

調査対象：第 11 次調査の回収標本 2,350 名。

調査方法：往復ハガキによる郵送調査法。ただし、督促は行わなかった。実査は社団法人 中央調査社に委託した。

調査時期：2003 年 11 月 12 日から 14 日にかけて投函し、2004 年 1 月 22 日までの到着分を返送³⁾とした。

表 2 第 11 次全国調査とその事後調査の回収状況

	第 11 次全国調査			事後調査		
	計画	回収	回収率	発送	返送	返送率
全体	4,193	2,350	56.0%	2,350	1,169	49.7%
K 型	2,088	1,192	57.1%	1,192	579	48.6%
M 型	2,105	1,158	55.0%	1,158	590	50.9%

²⁾ 本稿では計算の都合上、沖縄県を 20 万人未満の市部に合併し、4 層として扱った。

³⁾ 本稿では、紛らわしさを避けるため、事後調査における回収を返送と呼ぶことがある。

3. 傾向スコアを利用した方法

Y_i を対象者*i*が回収できるか否かを示すダミー変数（回収できれば $Y_i=1$ ，そうでなければ $Y_i=0$ ）とし， X_{i1}, \dots, X_{iH} を*H*個の補助変数とすると， $Y_i=1$ となる条件付確率 θ_i （Rosenbaum and Rubin (1983) や Rosenbaum (2002) はこれを傾向スコアと呼んでいる）

$$\theta_i = \Pr(Y_i=1 | X_{i1}, \dots, X_{iH}) \quad (5)$$

を利用し，

$$w_i = N/n \cdot 1/\theta_i \propto 1/\theta_i \quad (6)$$

とする。ただし，真の θ_i は知られていないため，本稿では，ロジスティック回帰により推定した $\hat{\theta}_i$ を用いる。最近では，Web 調査の補正に傾向スコアが利用されることがある（星野・繁樹，2004）。

4. キャリブレーションによる方法

H 個の補助変数全てに関して，キャリブレーション方程式

$$\sum_R w_i X_{ih} = \sum_U X_{ih} \quad (7)$$

あるいは

$$\sum_R w_i X_{ih} = N/n \sum_S X_{ih} \quad (8)$$

が成り立つという条件の下で，

$$\sum_R G(w_i, d_i) \quad (9)$$

を最小とする w_i を求める（Deville and Särndal, 1992; Deville, Särndal and Sautory, 1993; Särndal and Lundström, 2005）。ただし， $G(\cdot)$ は距離関数であり，本稿では以下の二つを取り上げる⁵⁾。

(a) linear : $G(w_i, d_i) = (w_i - d_i)^2 / 2d_i$

(b) log : $G(w_i, d_i) = w_i \log(w_i/d_i) - w_i + d_i$

初期ウェイト d_i は，通常は $d_i = N/n$ とし，4.4 節の二段推定では，一段目で得られたウェイトを d_i とする。また，本稿で用いる補助変数は全てカテゴリーカルであるため，カテゴリごとにキャリブレーション方程式（7）式あるいは（8）式を考えることになる。なお，距離関数が linear の場合，（2）式はいわゆる回帰推定量（Cochran, 1977; Särndal, Swensson and Wretman, 1992）となることが知られている。

以下の結果では，見やすさのため，得られたウェイト w_i に対して， $c \sum_R w_i = n^*$ を満たす c を乗じた cw_i を示している。

4. 事後調査データを用いた検討

この章では，事後調査の返送者データをウェイト集計した結果と，第11次調査の全回収標本の回答分布とを比較する。つまり，第11次調査全体の標本比率が「母集団における比率 p_{kc} 」となる。事後調査の未返送者を調査不能者として扱う際の注意点は土屋（2005）を，さらに一

⁵⁾ linear では，ウェイト w_i の最小値が負となることがある。そのため， $w_i \geq 10^{-4}$ という制約を設けた bounded linear についても検討したが，その結果は，linear とほとんど変わらないため，本稿では割愛する。

般に、パネルデータを利用して調査不能を検討する際の注意点は Groves and Couper (1998) を参照されたい。補正効果を測る指標としては、調査票ごとに

$$D = 100 \times \frac{1}{Q} \sum_{k=1}^Q \left\{ \frac{1}{C_k} \sum_{c=1}^{C_k} (p_{kc} - \hat{p}_{kc,w})^2 \right\}^{1/2} \quad (10)$$

を用いる。\$C_k\$ は質問項目 \$k\$ のカテゴリ数、\$Q\$ は質問項目数であり、K 型が \$Q=49\$ 項目、M 型が \$Q=63\$ 項目である。K 型と M 型を併合した \$Q=112\$ とせず、調査票ごとに \$D\$ を求めるのは、二例に限られてはいても実例数を増やし、より一般化可能な知見を得るためである。

4.1 回収率を利用した方法

表 3 の一行目は、(10) 式の \$\hat{p}_{kc,w}\$ を返送者の標本比率とした場合の \$D\$ である。同様に、二～三行目は、PSU あるいは層ごとに回収率の逆数をウェイトとする方法の \$D\$ である。カッコ内の数値は、標本比率の場合の \$D\$ に対するパーセントである。回収率の逆数をウェイトとする方法では、標本比率と比較して \$D\$ の値があまり小さくなっておらず、補正効果はほとんど見られないと言ってよい。

4.2 人口統計学的属性変数のみを補助変数とする場合

次に、表 4 の 8 種類の人口統計学的属性変数のいくつかを補助変数としてウェイトを求める (表 6 も参照)。また、任意の 2 種類の属性変数 A と B をクロスさせた変数 (これをクロス変数と呼び、A×B と表す) についても考える。クロス変数は全部で 28 個となる。3 種類以上の属性変数のクロスは、標本サイズが限られており、組み合わせの数が膨大となるため用いない。また便宜上、表 4 に示す属性変数を「種類」と数え、(5) 式や (7) 式、(8) 式の補助変数の数 \$H\$ を「個」で数える。例えば、性と年齢層の 2 種類を使用する場合、両変数をクロスさせず \$H=2\$ 個とする場合と、クロスさせて \$H=1\$ 個とする場合があり得る。なお、表 4 の属性変数は、(10) 式の \$Q\$ 個の質問項目には含まれない。前節の PSU あるいは層も補助変数の候補であるが、本稿では用いない。PSU の数は 300 と多く、層をより細かくしたものが表 4 の区市郡だからである。

4.2.1 補助変数が 1 個の場合

1 個のカテゴリカルな補助変数を利用する場合、傾向スコアを利用する方法とキャリブレーションによる方法は、いずれも事後層化推定量に等しい。表 4 の左半分には、各属性変数で事後層化したときの \$D\$ を示した。M 型の年齢層を除き、いずれの属性変数で事後層化した場合も、標本比率と大きな違いはない。地方や、K 型の住居形態、M 型の性では、標本比率よりも \$D\$ は大きい。

表 4 には、事後調査への返送・未返送と各属性変数とのクロス集計表の \$\chi^2\$ 値、および返送・未返送を基準変数、8 個全ての属性変数を同時に説明変数としたロジスティック回帰の結果も示した。M 型の年齢層のように、返送・未返送と関連した変数は \$D\$ を小さくする。しかし、K 型の年齢層のように強い関連は必ずしも小さな \$D\$ を保証せず、M 型の地方のように標本比率より \$D\$ を大きくすることもある。つまり、返送・未返送との関連の強さは、補助変数選択の

表 3 母集団との分布間距離 \$D\$

方法	K 型	M 型
標本比率	1.689872 (100.0%)	1.346058 (100.0%)
回収率を利用した方法 (PSU ごと)	1.732815 (102.5%)	1.271918 (94.5%)
回収率を利用した方法 (層ごと)	1.669406 (98.8%)	1.340621 (99.6%)

表4 人口統計学的属性変数

属性変数	母集団との分布間距離 D (標本比率の D に対する比)			返送・未返送とのクロス集計		返送・未返送を 基準とした回帰	
	K 型	M 型	自由度	χ^2 値		Wald χ^2 値	
				K 型	M 型	K 型	M 型
性	1.687 (99.8%)	1.346 (100.0%)	1	1.39	0.01	0.41	0.81
年齢層	1.681 (99.5%)	1.200 (89.1%)	5	29.86***	41.56***	24.07***	38.11***
区市郡	1.673 (99.0%)	1.341 (99.6%)	6	4.50*	6.09	6.64	6.09
地方	1.699 (100.5%)	1.363 (101.2%)	8	11.49	19.09**	14.19*	18.41**
世帯人数	1.662 (98.3%)	1.301 (96.7%)	5	11.60**	12.53**	6.46	8.29
住居形態	1.692 (100.1%)	1.314 (97.6%)	1	5.36**	4.08**	3.76*	1.56
学歴	1.665 (98.5%)	1.341 (99.6%)	2	1.10	2.31	11.16***	16.43***
就業状態	1.678 (99.3%)	1.343 (99.8%)	1	6.80***	0.06	1.81	0.44

*: $p < .10$ ** : $p < .05$ *** : $p < .01$

一つの基準ではあるが、絶対的な基準ではない。

図1の左右それぞれの図の左端の箱ヒゲ図は、クロス変数も含め、1個の補助変数を用いたときの、合計 $8+28=36$ 通りの D の分布である。縦軸は、標本比率の D に対するパーセントである。適切なクロス変数を用いると、クロスしない場合より D を小さくできる (K型, M型ともに年齢層×学歴で、それぞれ92.7%と83.2%) 一方、不適切なクロス変数は D を大きくしてしまう (K型は年齢層×区市郡で102.8%, M型は地方×住居形態で102.0%)。

4.2.2 傾向スコアを利用した方法とキャリブレーションによる方法の比較

$H \geq 2$ 個の補助変数を同時に用いる場合、クロス変数も含めた補助変数の組み合わせは全部で7,156通りある。図1の左2番目から7番目に、クロス変数を含まない場合 ($2^8 - 1 - 8 = 247$ 通り) と含む場合 (7,156 - 247 = 6,909通り) とに分けて D の分布を示した。×印は、8種類8個の場合の D である。1個よりも複数個の方が、クロス変数を含まないより含む方が、 D は小さい傾向にある。ただしクロス変数によっては、逆に D を大きくする。

特に図の「箱」に着目すると、キャリブレーションでは、linear と log の間に顕著な違いはない。一方、傾向スコアは、キャリブレーションより D が大きいようである。確認のため、図2に、傾向スコアの D とキャリブレーション (linear) の D との差を縦軸として7,156通りの組み合わせをプロットした。傾向スコアの方が D が大きい組み合わせは、K型が4,775通り、M型が5,281通りである。

傾向スコアの D が大きくなりやすい主な理由は、(6)式で確率の逆数をウェイトとすることにある。つまり、少数の回答者に大きなウェイトが置かれ、推定値が不安定になるのである (Little, 1986; Little and Rubin, 2002)。図3は、補助変数が8種類8個のときの、傾向スコアあるいはキャリブレーションによるウェイトのヒストグラムである。明らかに、傾向スコアによるウェイト分布は右裾が重い。

図2の横軸は、傾向スコアとキャリブレーションの間での、ウェイトの歪度の差である。歪度 S は、

$$S = \left\{ \frac{1}{n^*} \sum_R (w_i - \bar{w})^3 \right\} / \left\{ \frac{1}{n^* - 1} \sum_R (w_i - \bar{w})^2 \right\}^{3/2}, \quad \bar{w} = \frac{1}{n^*} \sum_R w_i \quad (11)$$

により求めた。傾向スコアによるウェイトは、ほとんどの場合、キャリブレーションによるウェイトよりも歪度が大きいこと、歪度の差が大きい場合に D の差が広がること、が分かる。

以上の結果を踏まえ、以下では、ウェイトは主にキャリブレーション (linear) により求め

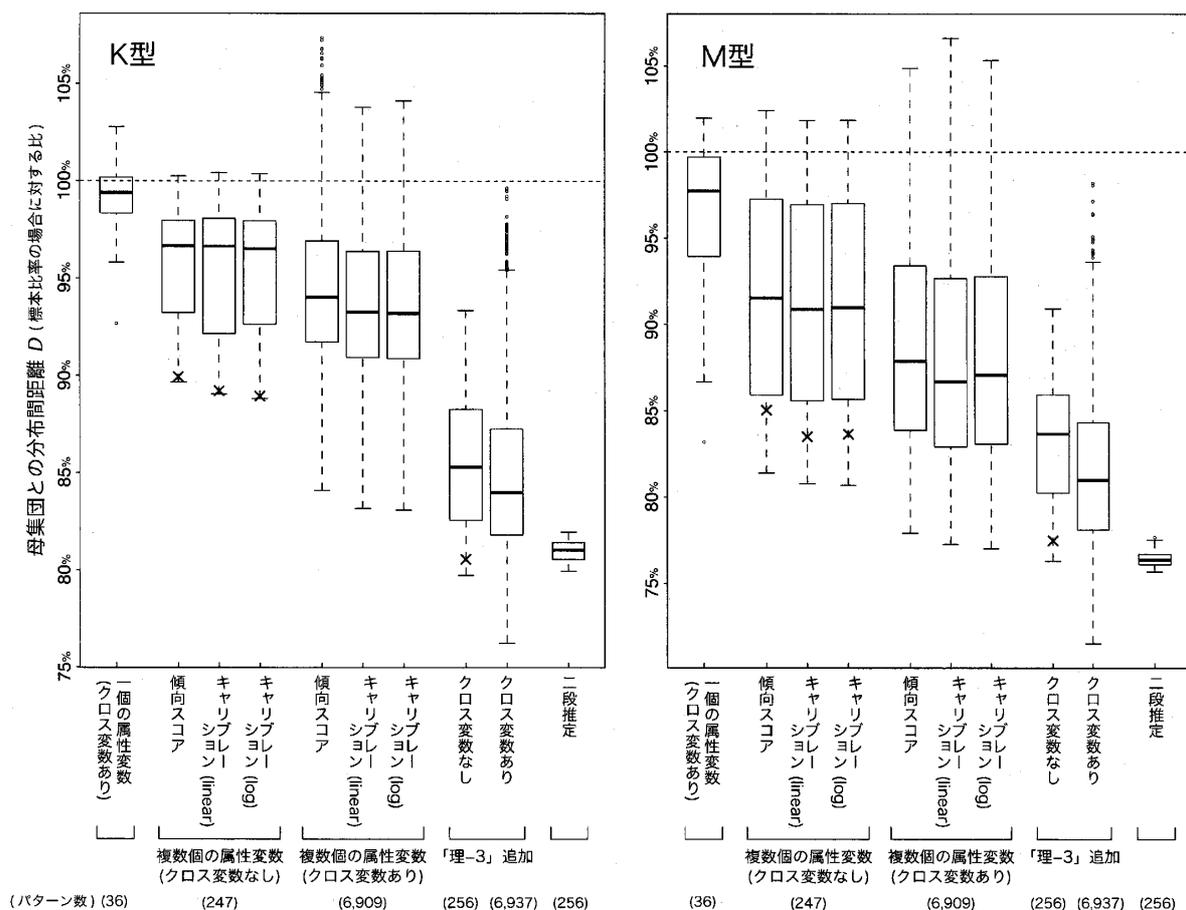


図1 母集団との分布間距離 D の補正方法間の比較

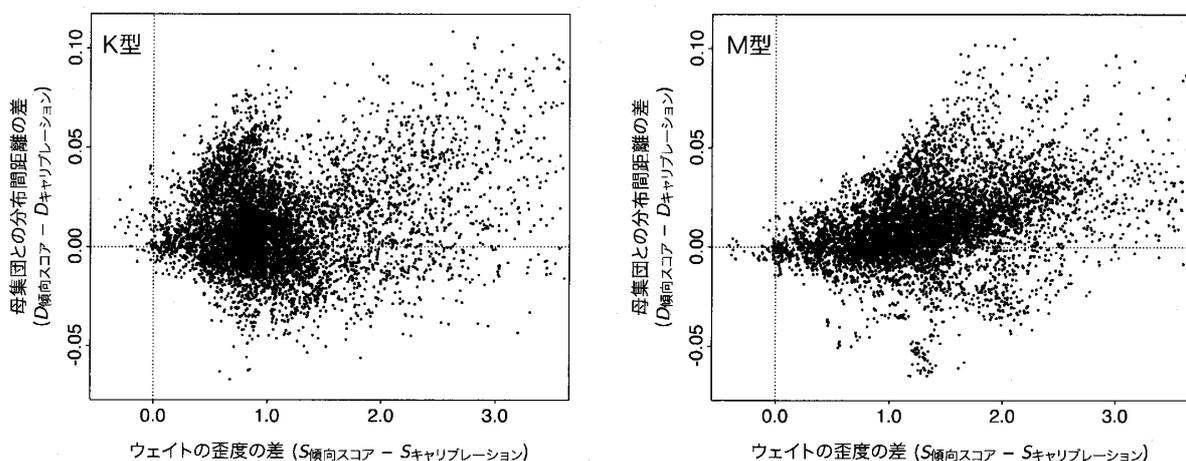


図2 傾向スコアとキャリブレーションとの間での、ウェイトの歪度 S の差と母集団との分布間距離 D の差

4.2.3 どの補助変数の組み合わせを用いればよいのか

最小の D は、必ずしも全ての補助変数を使う場合に得られるとは限らない。図4には、補助変数の種類と個数ごとに D の分布を示した。図の下部の点線は、8種類8個の場合の D である。一般に、補助変数の種類が多いほど D は小さい。しかし、8種類8個と同程度の D を得るには、K型では4種類3個、M型では2種類1個の適切な補助変数だけで十分である。

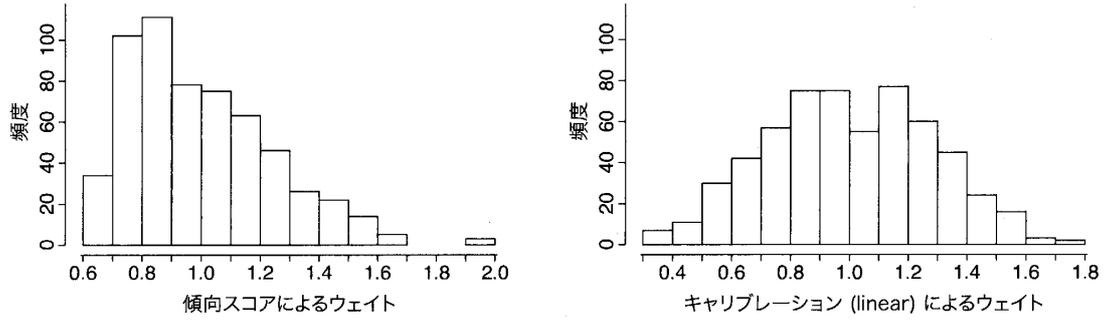


図3 傾向スコアによるウェイトの分布とキャリブレーションによるウェイトの分布の例

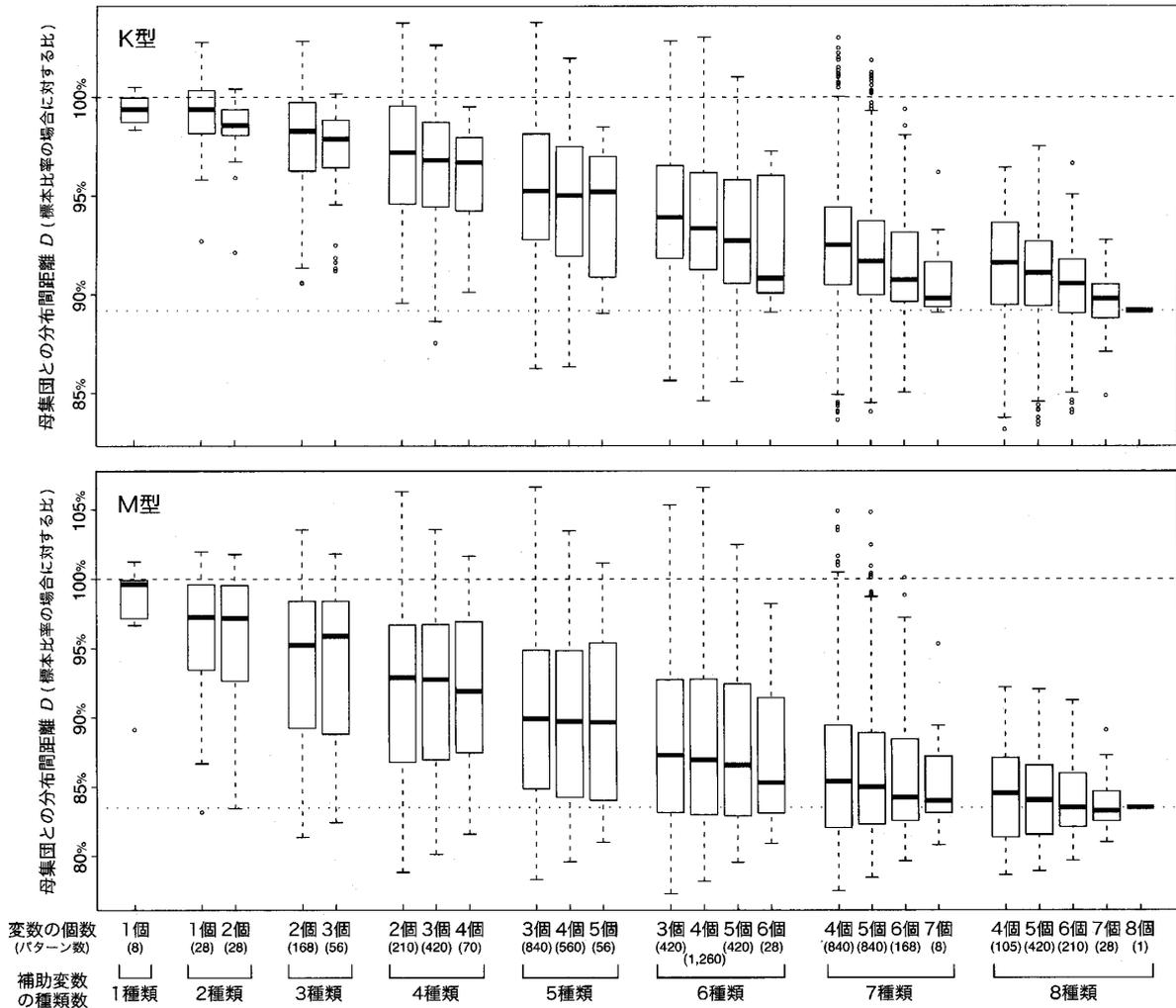


図4 キャリブレーションに用いる補助変数の種類・個数と母集団との分布間距離 D

さらに、K型では8種類4個でDは最小となるが、M型では6種類3個（性×世帯人数、年齢層×住居形態、学歴×就業状態）でDは最小となる。

Särndal and Lundström (2005) は、適切な補助変数を選ぶ指標の一つとして、ウェイトの分散

$$V = \frac{1}{n-1} \sum_R (w_i - \bar{w})^2 \quad (12)$$

の利用を提案している (p. 118). また, 項目への回答を補助変数にウェイトつきで回帰させたときの決定係数の利用も提案している (p. 122). 本稿では, $R_{kc,w}^2$ を項目 k の選択肢 c を基準変数としたときの決定係数とし,

$$\bar{R}^2 = \frac{1}{Q} \sum_{k=1}^Q \frac{1}{C_k} \sum_{c=1}^{C_k} R_{kc,w}^2 \quad (13)$$

を求める. さらに, 返送者の標本比率 $\hat{p}_{kc,w=N/n}$ と $\hat{p}_{kc,w}$ との間の距離

$$D^{\text{標}} = 100 \times \frac{1}{Q} \sum_{k=1}^Q \left\{ \frac{1}{C_k} \sum_{c=1}^{C_k} (\hat{p}_{kc,w=N/n} - \hat{p}_{kc,w})^2 \right\}^{1/2} \quad (14)$$

も考えることとする.

図5の上段二つの図には, V と D を用いて, 7,192 通りの補助変数の組み合わせをプロットした. 同様に, 中段では \bar{R}^2 と D , 下段では $D^{\text{標}}$ と D を用いた. D を小さくするには, V や \bar{R}^2 , $D^{\text{標}}$ がある程度大きい必要がある. しかしこれらの指標は絶対的な基準ではない. 結果は割愛するが, 他の指標として, 補助変数のカテゴリ数, ウェイトの歪度・尖度・レンジのい

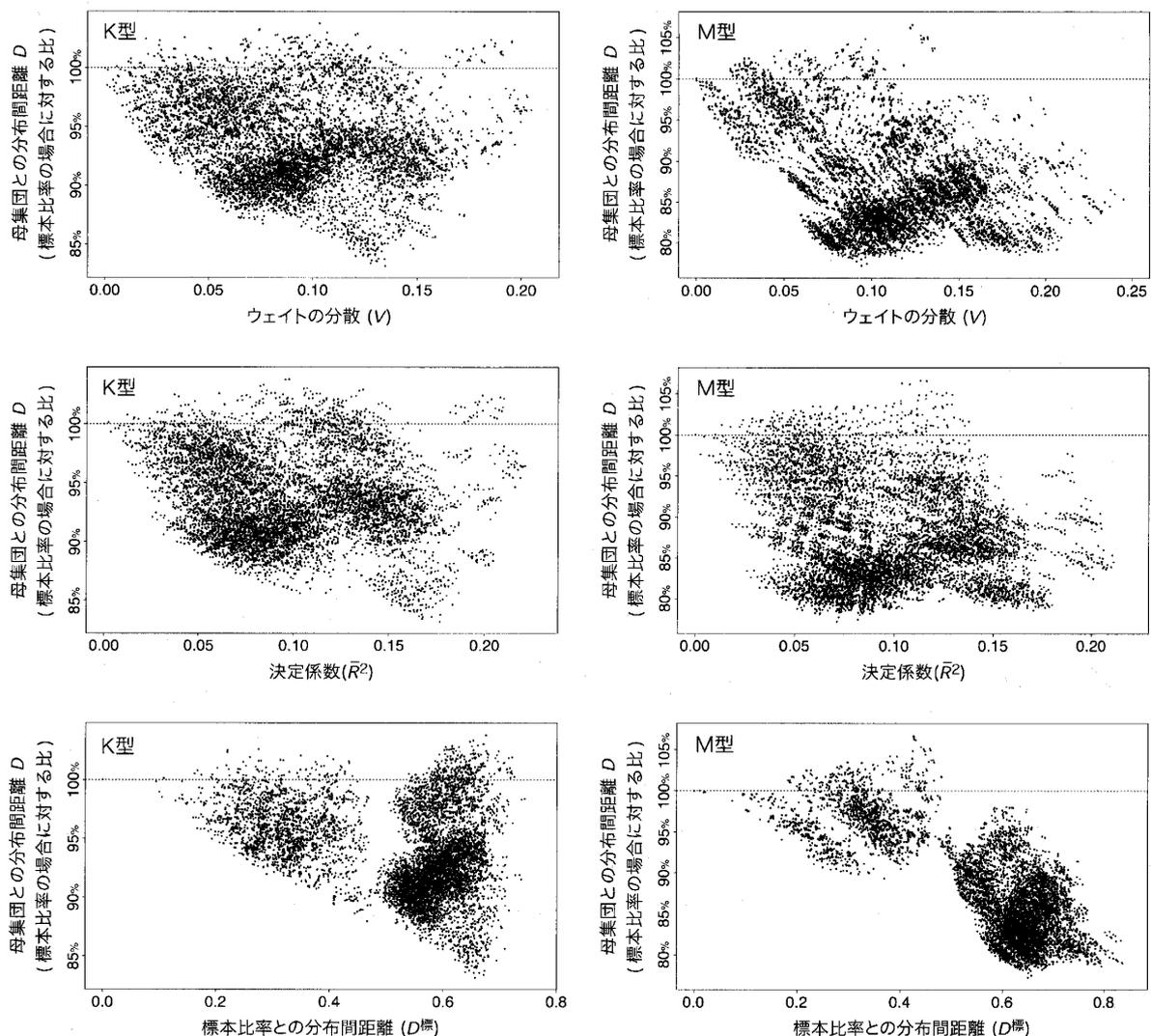


図5 ウェイトの分散 V , 平均決定係数 \bar{R}^2 , 標本比率との分布間距離 $D^{\text{標}}$ と母集団との分布間距離 D

ずれも絶対的な基準ではなかった。以上の指標は、任意の二つの属性変数をクロスさせるべきか否かの判断基準にもなり得なかったし、8種類全ての属性変数を用いる764通りの組み合わせに限定した場合でも、最小の D を選ぶ基準とはならなかった。

4.3 「調査への指向性」変数を補助変数とする場合

ところで、回収標本の中でも、調査協力の得やすさは必ずしも一様ではない。そして調査不能標本は、計画標本の中で最も協力が得にくい（得られない）対象者群である。つまり、調査協力の得やすさ、あるいは対象者の調査への指向性の強さを示す変数（本稿では広い意味を含めるため、これを指向性変数と呼ぶ）の分布は、母集団に比べ、回収標本では指向性が強い方へ偏る。そこで、指向性変数を補助変数とすれば、調査不能バイアスを低減できる可能性がある。

例えば Politz and Simmons (1949) は、不在によるバイアス補正のため、回答者に在宅確率を尋ね、その逆数をウェイトとすることを提案している。Lin and Schaeffer (1995) は、補助変数として架電回数や調査依頼時の拒否の程度を試みているし、Proctor (1996) や林 (2004) は、郵送法の督促回数や調査票返送日を補助変数とする方法を紹介している。本稿では、不能理由の大半を拒否が占める現状に鑑み、回答者の態度に注目して数種類の変数を比較し、指向性変数が持つべき性質を検討する。

また、母集団あるいは計画標本における指向性変数の分布は、客観的な資料としては存在し得ない。そこで、指向性変数として以下の二値変数を用いることを提案する。

$$X_{ih} = \begin{cases} 1 & \text{強指向性回答者 } i \in R \\ 0 & \text{弱指向性回答者 } i \in R \\ 0 & \text{調査不能者 } i \in S - R \end{cases} \quad (15)$$

回収標本の指向性の程度をさらに細かく分類できることもあろうが、調査不能者の分類は多くの場合不可能である。Holt and Elliot (1991) は、不能標本を4回目のアタックで回収できた標本で代替することを提案しているが、これは、弱指向性回答者を4回目での回収標本とした指向性変数で事後層化することに相当する。

4.3.1 指向性変数を追加した補正

この節では、指向性変数として事後調査の「調査への協力理由」を用いる。この質問への回答は、協力理由を表すと同時に、調査に対して積極的・消極的という態度を表すとも考えられる。そして郵送調査である事後調査では、24通の不達を除く1,157通の未返送は、未返送者自らの、返送しないという「調査に対する消極的態度」の表明であると言える。そこで、「1. 役に立ちたいと思ったから」「2. 内容がおもしろそうだったから」「3. 文部科学省の研究所が行っている調査だから」のいずれかを選んだ積極的な返送者を、(15)式の強指向性回答者とする（この指向性変数を、選択肢の番号から「理-3」と呼ぶ）。

図1の左右それぞれの右から3番目の箱ヒゲ図は、「理-3」と、クロス変数を除く属性変数を用いた256通りの D の分布である。ウェイトはキャリブレーション (linear) により求めた。属性変数のみを用いる左から3番目の箱ヒゲ図に比べ、明らかに D は小さい。さらに、最大の D でも標本比率を超えない。×印は補助変数が9個の場合の D であり、属性変数8個だけの×印に比べ、K型は89.2%が80.6%へ、M型へ83.5%が77.5%へと小さくなっている。この D は、属性変数だけで最適なクロス変数を用いたときの D に匹敵するか、さらに小さい。

右から2番目の箱ヒゲ図は、クロス変数を含む属性変数と「理-3」による6,937通り（属性変数と「理-3」とのクロスはない）の D の分布である。適切なクロス変数は D をさらに小さくするが、不適切なクロス変数は逆に D を大きくすることは、「理-3」を用いない場合と同

様である。

なお詳細は割愛するが、「理-3」を用いる場合でも、傾向スコアはキャリブレーションより D が大きくなりやすかった。具体的には、クロス変数を含む 7,192 通りのうち、K 型では 5,027 通り、M 型では 4,420 通りで、傾向スコアの方が D が大きい。

「理-3」の補正効果が大きい理由は、ウェイト分布を見るとよく理解できる。図 6 の左は、属性変数 8 個の場合のウェイト分布であり、右は「理-3」を加えた 9 個の場合である。グレー部分は強指向性回答者のウェイトを示す。事後調査を分析した土屋 (2005) によれば、積極協力者は未返送者とは回答が大きく異なる (p. 50)。属性変数だけの場合、調査不能者とは異なる性質を持つ強指向性回答者に対しても大きなウェイトが与えられてしまうのに対し、指向性変数を用いることでウェイトが抑えられ、調査不能バイアスが確実に低減されたのである。

4.3.2 どのような変数が指向性変数として適切か

この節では、3 種類の指標を取り上げ、指向性変数が持つべき性質を検討する。第一の指標は、4.3.1 節で用いた「協力理由」である。ただし、「理-3」の他にも、選択肢 1 の返送者のみを強指向性回答者とする変数（「理-1」と略）、選択肢 1 あるいは 2 のいずれかの場合とする変数（「理-2」と略）など、順に「理-6」まで考える。

第二の指標は、第 11 次調査時の「調査員判断」である。調査員は、面接調査終了後に、回答者の反応を「1. 最初から好意的」「2. だんだん好意的になった」「3. だんだん迷惑そうになった」「4. 最初から迷惑そう」の中から一つ選んでいる。選択肢 1 の返送者を強指向性回答者とする変数（「判-1」と略）、選択肢 1 あるいは 3 のいずれかの場合とする変数（「判-2」と略）、選択肢 4 以外の場合とする変数（「判-3」と略）の三つを考える。

第三の指標は、質問項目中の「D.K. 有無」である。Rizzo, Kalton and Brick (1996) によれば、パネル調査における調査継続の成否は、欠損項目の有無 (items imputed) が強く関係している。また土屋 (2005) は、事後調査における未返送者の D.K. の多さを指摘している。つまり、D.K. の多さは調査指向性の弱さを表すと考えられる。そこで、D.K. が全くない場合を強指向性回答者とする変数（「D-1」と略）、返送者のみの標本比率で D.K. が 3% を超える項目全てにおいて D.K. がない場合を強指向性回答者とする変数（「D-2」と略）の二つを考える。「D-2」は、D.K. があるとしても、それが稀なものであれば許容して強指向性回答者とするものである。

図 7 には、クロス変数を除く属性変数に指向性変数を加えた、各 256 通りの D の分布を示した。ただし、最も左は属性変数のみの場合である。図の「強指向性の割合」とは、返送者の中の強指向性回答者の割合である。「理-1」から「理-3」は補正に有効である一方、「理-4」か

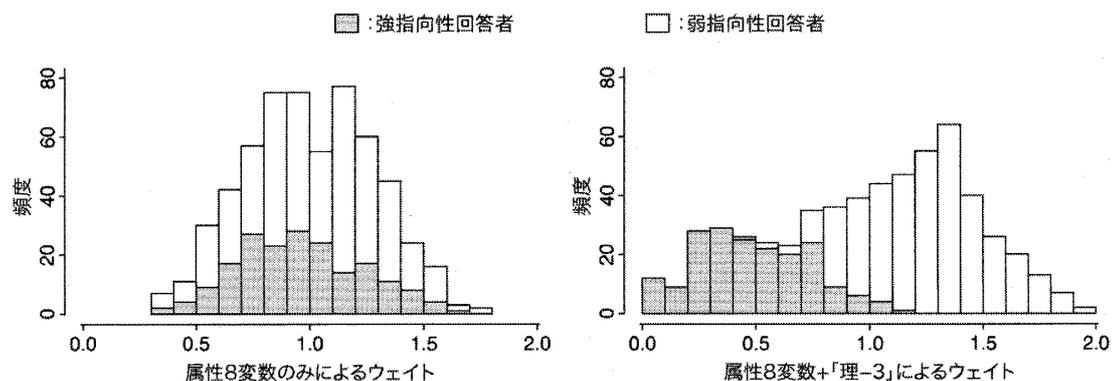


図 6 属性 8 変数のみによるウェイトの分布と「理-3」を加えたウェイトの分布 (K 型)

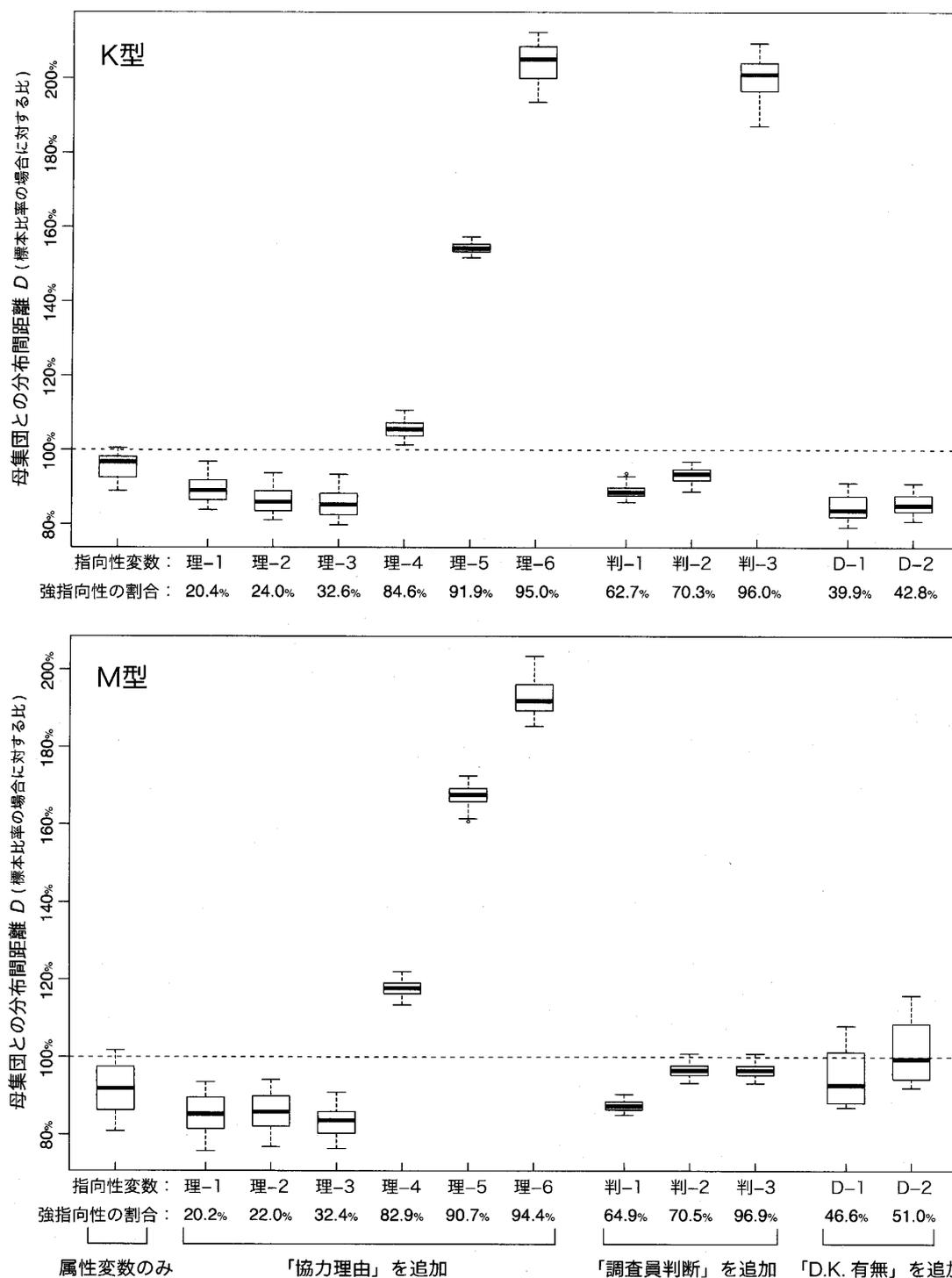


図7 母集団との分布間距離 D の指向性変数間の比較

ら「理-6」や K 型における「判-3」、M 型における「D.K. 有無」は、標本比率よりもバイアスを拡大している。

この結果は、強指向性回答者を、回答者の中でも指向性が強い一部の者に限るのがよいことを示唆する。「理-3」の強指向性回答者の割合が3割強なのに対し、「理-4」は8割を超え、「判-1」では6割以上である。図8は、属性変数8個に、「理-1」「理-4」「判-1」「D-1」のいずれか一つを加えた補助変数9個の場合のウェイト分布である。「理-4」や「判-1」など強指向性回答者の割合が高い指向性変数では、ウェイト分布が二峰化し、極言すれば少数の弱指向

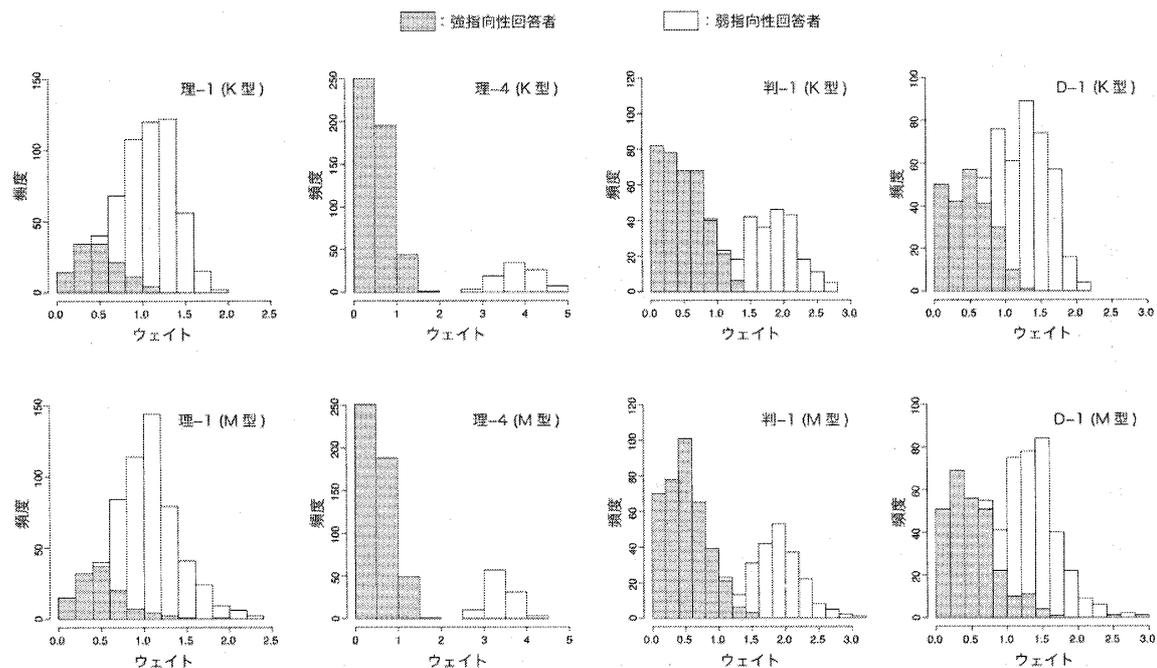


図8 各指向性変数を用いた場合のウェイト分布

性回答者の回答分布が推定値となってしまう。

つまり、指向性変数の作成にあたっては、調査不能者に似ている弱指向性回答者を見出し、その回答分布で調査不能者を代替する (Lynn, Clarke, Martin and Sturgis, 2002), という視点ではなく、調査不能者とは全く異なる強指向性回答者を見出し、そのウェイトを小さくする、という視点が必要である。この両者は、一見同じようだが、指向性変数の焦点位置は異なる。前者は、指向性の弱さをより細かく分類しようとするのに対し、後者は、指向性の強さを細かく識別しようとする。第11次調査の「調査員判断」は前者であり、そのため選択肢四つのうち「だんだん好意的になった」「だんだん迷惑そうになった」「最初から迷惑そう」の三つを用いて、指向性の弱さを識別している。むしろ、返送者の6割以上を占める「最初から好意的」の程度をより細かく識別すべきということである。

なお、K型とM型の間で「D.K.有無」の効果が異なった理由は、5.2節で述べるとおり、両調査票の間で質問・選択肢の性格が異なるためであろう。調査指向性の指標として「D.K.有無」の無批判な使用は危険と言える。

4.4 二段推定

指向性変数を、母集団や計画標本における分布が事実として知られている属性変数と対等に扱い、キャリブレーションを行うことには、心理的な抵抗感があるかもしれない。むしろ、回収・不能を規定する一要因としてモデル化し、傾向スコアの算出に用いる方が指向性変数には馴染むという考え方もあろう。ただし、傾向スコアはキャリブレーションに劣ることが多い。そこで、まず指向性変数といくつかの属性変数を用いて傾向スコアによる補正を行い、その後属性変数だけでキャリブレーションするという二段推定を検討する。二段階で補正を行う例は、Kalton and Kasprzyk (1986) にも見られる。

図1の右端の箱ヒゲ図は、まず、クロス変数を除く属性変数と「理-3」を用いて傾向スコアにより第一段のウェイト $w_i^{(1)}$ を求め、次に $w_i^{(1)}$ を(9)式の d_i とし8個全ての属性変数を用いてキャリブレーションを行った結果である。二つ左隣の×印が、「理-3」と属性変数8個を用いた(一段の)キャリブレーション結果であり、二段推定の結果はこれとほとんど変わらな

い。

図9の左は、「理-3」と属性変数8個を用いた傾向スコアによるウェイト分布である。図6右のキャリブレーションによるウェイト分布と比べ、強指向性回答者のウェイトの分散がかなり小さい。図9の右は二段推定の結果であるが、一段目で既に属性変数8個を用いているため、左右の図の間に大きな違いは見られない。

つまり、傾向スコアによるウェイトでは強指向性回答者が極端に抑制されてしまい、二段推定によってその後キャリブレーションを行っても回復は難しい。傾向スコアを利用する場合には、キャリブレーションの場合以上に指向性変数を慎重に設定する必要がある。

4.5 事後調査データを用いた分析のまとめ

ここでひとまず、事後調査データを利用した検討結果を、箇条書きでまとめる。

- 基本的に、標本比率のままではなく、補正は行う方がよい。ウェイトは、傾向スコアを利用するよりもキャリブレーションにより求める方がよいことが多い。
- 一般に、補助変数はなるべく多い方がよい。クロス変数を用いるとバイアスを軽減できることが多い一方、増大させてしまうこともある。クロス変数も含めた、適切な補助変数の組み合わせを選ぶ絶対的な基準を見出すことはできなかった。
- キャリブレーションには指向性変数も併用すべきである。その趣旨は、調査不能者とは著しく異なる強指向性回答者に大きなウェイトを与えないようにすることである。得られたウェイトの分布は必ずチェックする必要がある。
- (一段の) キャリブレーションに比べ、二段推定が特に優れている、という結果は得られなかった。
- 指向性変数を用いて補正を行っても、調査不能によるバイアスは、標本比率の場合の(Dの測度で)8割程度残る。

5. 日本人の国民性第11次全国調査結果補正の試み

この章では、第11次調査結果全体の補正を試みる。補助変数とする属性変数は、これまでの8種類のうち世帯人数を除く7種類全てとし(表6)、クロス変数はいない。区市郡と住居形態に対しては(8)式、他の属性変数に対しては(7)式を用いて、キャリブレーション(linear)によりウェイトを求める。

5.1 指向性変数の選択

第11次調査の不能理由は、半数近くが拒否であり、次いで四分の一が一時不在である(表1)。

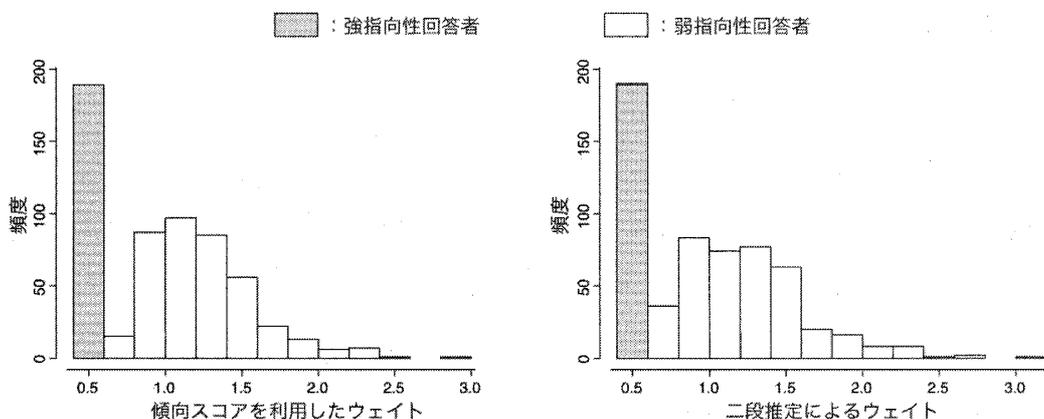


図9 傾向スコアを利用したウェイトと二段推定のウェイトの分布 (K型)

調査への指向性が強い回答者とは、これらの不能者の対極にある、調査に積極的で在宅していることが多い回答者となる。しかし調査時には、強指向性の識別に主眼が置かれておらず、適切な指向性変数の設定は難しい。本稿では便宜的に、「調査員判断」と事後調査の「協力理由」とを組み合わせ、強指向性回答者として以下の3パターンを検討する。

パターン1: 「協力理由」の選択肢1から3の回答者(「理-3」と同じ)

パターン2: 「協力理由」の選択肢1から4, かつ「調査員判断」の選択肢1の回答者

パターン3: 「協力理由」の選択肢1から3の回答者, または「協力理由」の選択肢4かつ「調査員判断」の選択肢1の回答者

図10は、属性変数7個と各指向性変数を用いたウェイト分布である。ウェイトの分散および回答者に占める強指向性回答者の割合も示した。強指向性回答者の割合が、パターン1では

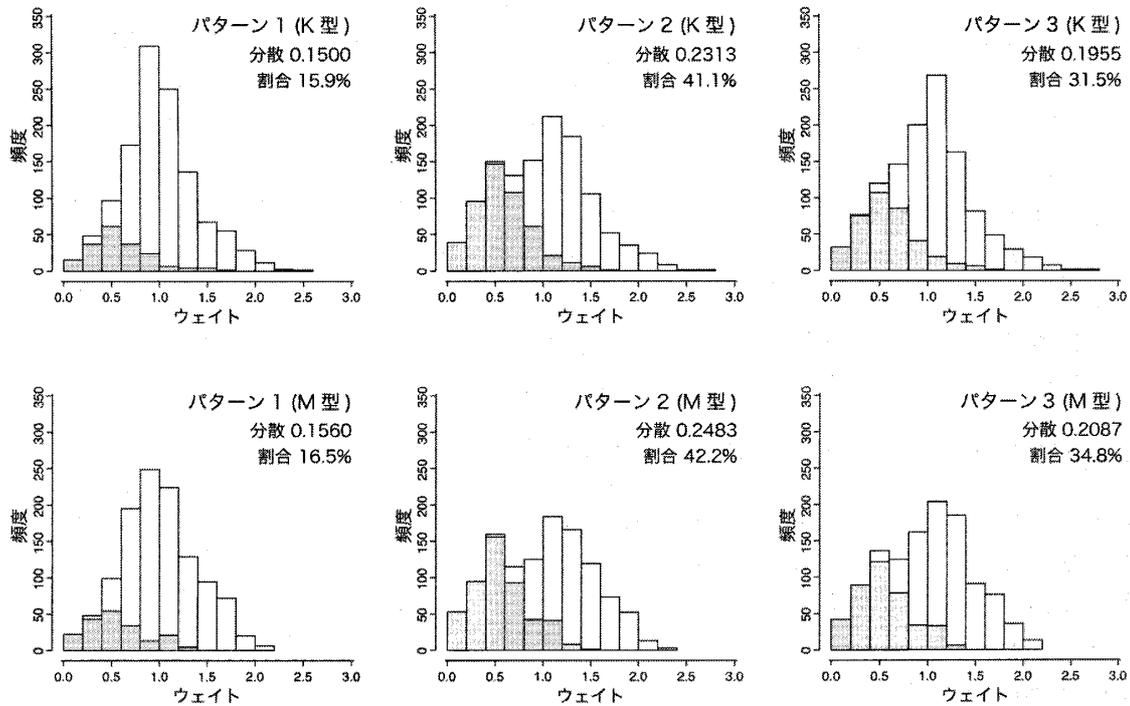


図10 各指向性変数を用いたウェイトの分布

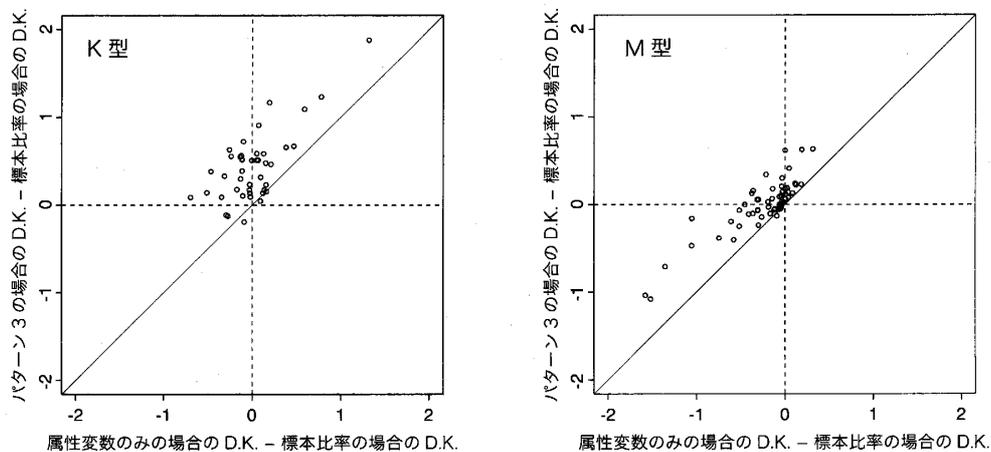


図11 D.K. のパーセントの比較

低すぎ、パターン2では高すぎるかもしれない。パターン3では、強指向性回答者と弱指向性回答者の分離がまだ見られず、三つの中では最も妥当と思われる。したがって、パターン3を採用する。

5.2 D.K. の比較

土屋 (2005) や Couper (1997), Triplett, Blair, Hamilton and Kang (1996) は、調査に対して消極的な回答者は D.K. が多いことを指摘している。つまり母集団では、標本比率よりも D.K. が大きいのではないかと考えられる。図 11 は、標本比率、属性変数のみによる補正結果、指向性変数も加えた補正結果、の三つの中で D.K. のパーセントを比較したものである。横軸は、属性変数による補正後の D.K. のパーセントから標本比率のそれを引いた値、縦軸は指向性変数も加えた補正值から標本比率を引いた値である。

K 型では、指向性変数を加えた補正は、標本比率よりも D.K. が大きいのに、属性変数のみによる補正では、必ずしも標本比率より D.K. が大きくなるとは限らない。ほとんど全ての点が対角線より上に位置し、指向性変数を加えることで、属性変数だけの場合より D.K. は大きくなる。

M 型でも、ほとんど全ての項目で、指向性変数を加えた補正值は、属性変数だけよりも D.K. は大きい。しかし、指向性変数を用いた場合でも、必ずしも標本比率より D.K. が大きくなるとは限らない。補正值の方が D.K. が小さい項目は、'#2.3i 仕事や職場に満足か' (-1.1 ポイント), '#9.12b 日本の「芸術」' (-1.0 ポイント), '#9.12 日本の「科学技術の水準」' (-0.7 ポイント) などである。このうち例えば #2.3i の D.K. は、回収率が低い 20 歳代で 6% (標本比率) なのに対し、回収率が高いが、この質問が無意味である人が多いであろう 70 歳以上では 43% であり、標本比率はむしろ高すぎる可能性がある。同様に #9.12b でも、20 歳代の D.K. は 7%、70 歳以上は 32% などと大差がついている。

また、K 型と M 型で様子が異なる原因の一つは、質問や選択肢の性格の違いにもあると考えられる。K 型には、二つの対立する意見の一方を選ぶという認知的に高負荷を要求する質問が多い。そのため調査に消極的な対象者が多いと、D.K. は増えやすい。一方、M 型には段階評定を求める質問が多い。この形式は認知的な負荷が高くなく、D.K. は発生しにくい。そのため、図 11 の M 型で、多くの点が原点付近に集まったのであろう。

以上をまとめると、指向性変数を加えた補正值は、多くの項目で他の二方法の集計値より D.K. が大きく、先述の予想に従えば三つの中で最も望ましい結果と考えられる。もちろん、D.K. が大きいほどよいわけではない。しかし、標本比率と比べ高々 1~2 ポイント程度の増加であり、D.K. が過大となった可能性は低いであろう。

5.3 回答選択肢の比較

付録の表 5 に、標本比率と補正值が大きく異なる調査項目を示す。具体的には、 \hat{p}_{kc} を標本比率、 $\hat{p}_{kc,w}$ を補正值とすると、

$$D_k^{\text{標}} = 100 \times \left\{ \frac{1}{C_k} \sum_{c=1}^{C_k} (\hat{p}_{kc} - \hat{p}_{kc,w})^2 \right\}^{1/2} \geq 1.4 \quad (16)$$

となる項目を示した⁶⁾。標本比率の 95% 信頼区間は、層化二段抽出の場合の \hat{p}_{kc} の標準誤差 (前田・中村, 2000) を s_{kc} とし、

⁶⁾ 1.4 という数値に意味はなく、適当な項目数にとどめるための値である。

$$L = \log \frac{\hat{p}_{kc}}{1 - \hat{p}_{kc}} - t_{.975, \nu} \frac{\hat{S}_{kc}}{\hat{p}_{kc}(1 - \hat{p}_{kc})} \quad (17)$$

$$U = \log \frac{\hat{p}_{kc}}{1 - \hat{p}_{kc}} + t_{.975, \nu} \frac{\hat{S}_{kc}}{\hat{p}_{kc}(1 - \hat{p}_{kc})} \quad (18)$$

とすると

$$\left(\frac{\exp(L)}{1 + \exp(L)}, \frac{\exp(U)}{1 + \exp(U)} \right) \quad (19)$$

により求めた。t分布の自由度は $\nu = 300 - 4 = 296$ であり、この95%信頼区間の外にある補正值は太字で示した。

補正効果が大きい項目を見ていくと、'#5.24 勤め先を変えるか'は、標本比率で「ながく勤めるのがよい」(49.1%)が「かわった方がよい」(46.5%)を上回るのに対し、指向性変数を追加した補正值では、「かわった方がよい」(51.1%)が過半数となる。標本比率でこの項目を時系列的に追うと、「かわった方がよい」は、1993年の41.4%から徐々に増え、2003年にやっと「ながく勤めるのがよい」に追いついている。しかし実際は、既に大きく逆転していると考えられるかもしれない。

'#2.11 好きなくらし方か人のためか'は、標本比率では10ポイント程度の差をつけて、「人のためになること」が「自分の好きなこと」より多いが、補正值では3.4ポイント差に縮まっており、伯仲している。同様のことは、'#5.6b つとめたい会社'についても言える。'#4.32 離婚すべきでないか'は、標本比率では「離婚すべきでない」(27.2%)と「合意さえあれば離婚」(26.9%)が拮抗している。しかし補正值は「合意さえあれば離婚」が4.7ポイント上回っており、離婚容認の流れが強まっているのではないかと考えられる。

参考として、本稿の冒頭で紹介した'#2.12c 人は信頼できるか'や'#2.12 他人のためか自分のためか'の補正結果を表5の最後に示す。「用心する」が59.2%から61.1%へ、「自分のことだけ」が50.8%から51.3%へと増えてはいるが、いずれも信頼区間に収まり、1978年からの四半世紀の変化の方向性を覆すほどの大きな増加ではない。

6. 考 察

4.5節と重複する部分もあるが、本稿の三つの目的に沿って、分析結果に検討を加えたい。

まず、ウェイトを求める方法としては、3章の四つの中ではキャリブレーションがよい。標本比率のままよりも、補正を行う方がよい、ということである。さらにこの結果は、補助変数の周辺分布のみ知られていることが多い現実の調査の補正にとって朗報と言える。傾向スコアによる方法は、その論理は明快である。しかし実際の適用場面では、調査不能者についても補助変数の値を必要とする。さらに、(6)式で逆数を用いることがウェイトの分散を拡大させ、逆に分散縮小を狙うキャリブレーションに安定性の面で劣る。本稿ではロジスティック回帰により $\hat{\theta}_i$ を求めたが、例えばAIDやCART (Breiman, Friedman, Olshen and Stone, 1984)などを用いても、(6)式による限り本質は変わらない。もちろん、以上の結論は推定値の比較に基づくものであり、推定値の分散に関する比較は今後の課題としたい。

次に、Kalton and Flores-Cervantes (2003)も述べているように、ウェイトを求める方法の選択よりも、補助変数の選択の方が、結果に与える影響は大きい。不適切な補助変数は標本比率よりバイアスを大きくする。補正に有効な補助変数は調査項目によって異なるであろうし、

どの変数が有効か見定める絶対的な基準も今のところ見当たらないようである。最も現実的な対処法は、可能な限り多数の補助変数を用いることである。有効な変数が含まれる可能性が高まり、仮に不適切な変数が含まれていてもその影響が相殺されると“期待”できるからである。

変数の数を増やすという意味では、客観的な指標ではなくとも、合理的な仮定に基づく指向性変数も補助変数として有効である。実際、事後調査における補正効果を実証したことは、本稿の成果の一つである。指向性変数には、さらに、調査不能者とは大きく異なる回答者のウェイトを低く抑えるという重要な役割がある。例えば、若年層や男性、都市部の回収率が低いのは現在に限ったことではない（鈴木, 1961; 坂元, 1975; DeMaio, 1980）。若年男性は調査に対する指向性が弱いのが普通であり、回収率が低いからといって指向性が強い若年男性回答者を過剰に拡大しないよう、指向性変数による抑制が必要である。

したがって、強指向性回答者を定める指標の選択が、補正効果の大きさを左右する。郵送法による事後調査では、不能者の多くが調査に消極的と考えられるため、積極・消極という態度に着目し、補正効果をあげた。第11次調査に限らず、面接調査の多くでは拒否による調査不能が増加しており（Synodinos and Yamada, 2000）、態度に着目することは有効であろう。同時に、拒否とは性質が異なる一時不在への対応も必要である。例えば Groves and Couper (1998) は、傾向スコアを利用し、一時不在と拒否を二段階で補正することを提案しているし、Iannacchione (2003) は、パネル調査結果を移転と拒否の二段階で補正している。不能理由に応じて複数の指向性変数を用意すべきか、あるいは一つの指向性変数にまとめるべきか、またどのような指標で指向性を測るべきか、今後の検討課題である。

操作的に定義される指向性変数は、当然、母集団における分布が知られていない。本稿では(15)式の利用を提案した。このことは、特に回収率が低い現状では必然的に、不能者に似た回答者を探すという従来の研究の方向から、不能者とは似ていない回答者を探すという方向への転換を要求する。似ているとは言っても、協力が得られた以上、本稿で言う弱指向性回答者と調査不能者とは本質的に異なる（O'Neil, 1979; Stinchcombe, Jones and Sheatsley, 1981; 杉山, 1984; Teitler, Reichman and Sprachman, 2003）と考えるのが自然である。その違いを捨象して $X_{ih}=0$ とまとめる一方、同じ協力者であっても二分割するのであるから、強指向性回答者は相当に偏った一部の標本に限定することになる。

第11次調査結果の補正では、必ずしも適切な指向性変数が用意できず、あくまで試算ではあるが、標本比率の95%信頼区間に収まらない補正值がいくつか得られた。さらに、数は少ないものの、結果の解釈が変わり得る項目も存在した。第11次調査の結果として、標本比率はかなり大きな不能バイアスを抱えているのではないかと危惧される。5年ごとという周期から言えば、次回の「日本人の国民性調査」は2008年の予定である。回収率の大幅な改善は期待できず、一層の悪化が懸念される。回収率が今ほど低くない時期には、補正結果は標本比率とさほど変わらないと考えられていたが（例えば、水野（1992）など）、今後は、標本抽出デザインの再考も含め、補正も視野に入れた準備が必要であろう。同時に、事後調査の例では、調査不能バイアスの改善が（ D の測度で）高々2割程度であったことを考えると、補正などの事後処理だけでは、バイアスの解消に楽観的な見通しは持てない。長期的な視野から、例えば学校教育や社会教育における「調査リテラシー」への取り組みなど、調査環境の抜本的な改善へ向けた努力が重要である。

謝辞

有益なコメントをしていただいた査読者の先生方と統計数理研究所 坂元慶行教授に感謝いたします。

参 考 文 献

- [1] Bethlehem, J. G. (2002). Weighting nonresponse adjustments based on auxiliary information. *Survey Nonresponse* (eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little), 275-287, John Wiley & Sons, New York.
- [2] Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees*. Chapman & Hall, New York.
- [3] Cochran, W. G. (1977). *Sampling Techniques*. 3rd ed. John Wiley & Sons, New York.
- [4] Couper, M. P. (1997). Survey introductions and data quality. *Public Opinion Quarterly*, **61**, 317-338.
- [5] DeMaio, T. J. (1980). Refusals: Who, where and why. *The Public Opinion Quarterly*, **44**, 223-233.
- [6] Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Jour. Amer. Statist. Assoc.*, **87**, 376-382.
- [7] Deville, J.-C., Särndal, C.-E. and Sautory, O. (1993). Generalized raking procedures in survey sampling. *Jour. Amer. Statist. Assoc.*, **88**, 1013-1020.
- [8] Ekholm, A. and Laaksonen, S. (1991). Weighting via response modeling in the Finnish household budget survey. *Journal of Official Statistics*, **7**, 325-337.
- [9] Gelman, A. and Carlin, J. B. (2002). Poststratification and weighting adjustment. *Survey Nonresponse* (eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little), 289-302, John Wiley & Sons, New York.
- [10] Groves, R. M., Cialdini, R. B. and Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, **56**, 475-495.
- [11] Groves, R. M. and Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*. John Wiley & Sons, New York.
- [12] 林英夫 (2004). 郵送調査法, 関西大学出版部.
- [13] Holt, D. and Elliot, D. (1991). Methods of weighting for unit non-response. *The Statistician*, **40**, 333-342.
- [14] 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, 第31巻第1号, 43-61.
- [15] Iannacchione, V. G. (2003). Sequential weight adjustment for location and cooperation propensity for the 1995 national survey of family growth. *Journal of Official Statistics*, **19**, 31-43.
- [16] Kalton, G. and Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, **19**, 81-97.
- [17] Kalton, G. and Kasprzyk, D. (1986). The treatment of missing survey data. *Survey Methodology*, **12**, 1-16.
- [18] Lin, I. F. and Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation. *The Public Opinion Quarterly*, **59**, 236-258.
- [19] Little, R. J. A. (1986). Survey nonresponse adjustments for estimates of means. *International Statistical Review*, **54**, 139-157.
- [20] Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. 2nd ed. John Wiley & Sons, New Jersey.
- [21] Lynn, P., Clarke, P., Martin, J. and Sturgis, P. (2002). The effects of extended interviewer efforts on nonresponse bias. *Survey Nonresponse* (eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little), 135-147, John Wiley & Sons, New York.
- [22] 前田忠彦・中村隆 (2000). 近年5回の国民性調査の標本設計と標本精度について 統計数理, 第48巻第1号, 147-178.
- [23] Meegama, N. and Blair, J. (1999). The effects of telephone introductions on cooperation: An experimental comparison. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 394-397.
- [24] 水野欽司 (1992). 基本項目 統計数理研究所国民性調査委員会 第5日本人の国民性一戦後昭和期総集一, 至誠堂, Pp. 45-65.
- [25] O'Neil, M. J. (1979). Estimating the nonresponse bias due to refusals in telephone surveys. *The Public Opinion Quarterly*, **43**, 218-232.
- [26] Politz, A. and Simmons, W. (1949). An attempt to get the "not at homes" into the sample without callbacks. *Journal of the American Statistical Association*, **44**, 9-16.
- [27] Proctor, C. (1996). Extrapolating on day returned to adjust for mail survey nonresponse. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 579-583.
- [28] Rizzo, L., Kalton, G. and Brick, J. M. (1996). A comparison of some weighting adjustment methods for panel nonresponse. *Survey Methodology*, **22**, 43-53.

- [29] Rosenbaum, P. R. (2002). *Observational Studies. Second Edition*. Springer-Verlag, New York.
- [30] 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.
- [31] 坂元慶行 (1975). 調査の種類と実施の概要 統計数理研究所国民性調査委員会 第3日本人の国民性, 至誠堂, Pp. 391-403.
- [32] 坂元慶行・中村隆・前田忠彦・土屋隆裕 (2004). 国民性の研究第11次全国調査 統計数理研究所研究レポート 92.
- [33] Särndal, C.-E. and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. John Wiley & Sons, Chichester.
- [34] Särndal, C.-E. and Swensson, B. (1987). A general review of estimation for two phases of selection with applications to two-phase sampling and nonresponse. *International Statistical Review*, **55**, 279-294.
- [35] Särndal, C.-E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer-Verlag, New York.
- [36] Stinchcombe, A. L., Jones, C. and Sheatsley, P. (1981). Nonresponse bias for attitude questions. *The Public Opinion Quarterly*, **45**, 359-375.
- [37] 杉山明子 (1984). 社会調査の基本, 朝倉書店, 東京.
- [38] 鈴木達三 (1961). 調査不能の分析と郵便調査による検討 統計数理研究所国民性調査委員会 日本人の国民性 至誠堂 Pp. 355-372.
- [39] Synodinos, N. E. and Yamada, S. (2000). Response rate trends in Japanese surveys. *International Journal of Public Opinion Research*, **12**, 48-72.
- [40] Teitler, J. O., Reichman, N. E. and Sprachman, S. (2003). Costs and benefits of improving response rates for a hard-to-reach population. *Public Opinion Quarterly*, **67**, 126-138.
- [41] Triplett, T., Blair, J., Hamilton, T. and Kang, Y. C. (1996). Initial cooperators vs. converted refusers: Are there response behavior differences? *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1038-1041.
- [42] 土屋隆裕 (2005). 調査不能者の特性に関する一考察—「日本人の国民性第11次全国調査」への協力理由に関する事後調査から— 統計数理, 第53巻第1号, 35-56.

表5 標本比率とキャリブレーションによる補正結果

選択肢	標本比率 (95% C.I.)	属性変 数のみ	指向性 変数	選択肢	標本比率 (95% C.I.)	属性変 数のみ	指向性 変数
<u>#2.3i 仕事や職場に満足か</u>				<u>#5.1b 親がキトクのとき</u>			
満足	12.6 (10.7-14.9)	11.8	11.3	故郷へ帰る	39.9 (36.8-43.2)	41.5	42.2
やや満足	36.4 (33.5-39.3)	36.2	35.7	会議に出る	51.3 (48.0-54.7)	50.3	49.0
やや不満	26.5 (24.0-29.2)	29.0	29.8	<u>#5.6b つとめたい会社</u>			
不満	5.9 (4.7-7.4)	6.2	5.9	給料が多い会社	43.6 (40.6-46.6)	47.0	47.6
<u>#2.10 幸福かためになることか</u>				家族的な雰囲気のある会社	52.8 (49.7-55.8)	49.9	49.0
しあわせにくらす	36.2 (33.2-39.2)	38.4	39.1	<u>#5.23 能力か功労か</u>			
世の中のためになる	58.9 (55.9-61.8)	56.7	55.3	能力重視	58.9 (56.1-61.6)	61.2	61.3
<u>#2.11 好きなくらし方か人のためか</u>				功労重視	28.4 (25.8-31.1)	26.7	26.5
自分の好きなこと	39.7 (36.6-42.9)	43.0	44.3	<u>#5.24 勤め先を変えるか</u>			
人のためになること	51.6 (48.4-54.9)	49.2	47.7	かわった方がよい	46.5 (43.7-49.4)	50.9	51.1
<u>#2.13 将来に備えるか楽しむか</u>				ながく勤めるのがよい	49.1 (46.3-52.0)	45.2	44.7
将来に備える方	65.8 (62.8-68.7)	62.8	61.7	<u>#6.2 男・女の生まれかわり</u>			
楽しむ方	29.9 (27.1-32.8)	33.0	33.9	男に	54.3 (51.4-57.1)	56.1	56.3
<u>#2.30e 不安感 失業</u>				女に	38.9 (36.3-41.7)	37.0	36.3
非常に感じる	19.8 (17.5-22.3)	19.2	18.3	<u>#7.1 人間らしさはへるか</u>			
かなり感じる	23.7 (21.4-26.3)	24.4	24.5	賛成	54.2 (50.8-57.5)	52.4	51.7
少しは感じる	33.2 (30.4-36.1)	35.5	35.8	いちがいにいいない	28.8 (25.9-31.8)	30.0	30.6
まったく感じない	20.6 (18.1-23.2)	18.5	18.8	反対	13.2 (11.3-15.3)	13.6	13.5
<u>#2.30f 不安感 戦争</u>				<u>#7.4b 国の繁栄と国民の生活</u>			
非常に感じる	28.3 (25.2-31.7)	26.4	24.9	よくなる	28.4 (25.9-31.1)	30.4	30.8
かなり感じる	23.1 (20.8-25.7)	23.8	24.0	よくなる	67.0 (64.1-69.8)	65.3	64.8
少しは感じる	32.9 (29.9-36.0)	34.6	35.4	<u>#7.5b* 公益と個人の権利</u>			
まったく感じない	14.6 (12.1-17.5)	14.3	14.7	個人の権利	38.5 (35.5-41.6)	41.1	40.9
<u>#2.30g 不安感 原子力施設の事故</u>				公共の利益	50.9 (47.7-54.1)	49.5	48.9
非常に感じる	26.3 (23.4-29.5)	24.3	23.4	<u>#8.6 選挙への関心</u>			
かなり感じる	23.4 (20.9-26.1)	23.9	23.0	なにをおいても投票	35.2 (32.0-38.4)	32.1	31.5
少しは感じる	33.3 (30.3-36.5)	35.2	36.2	なるべく投票	50.3 (47.2-53.3)	50.3	50.2
まったく感じない	14.8 (12.2-17.8)	14.5	15.2	あまり投票する気にならない	7.9 (6.4-9.6)	9.2	9.9
<u>#2.80c 病気 いろいろ</u>				ほとんど投票しない	6.1 (4.8-7.7)	7.5	7.6
あり	45.0 (41.7-48.3)	47.2	48.2	<u>#8.7i 支持政党</u>			
なし	54.6 (51.3-57.8)	52.5	51.5	自民党	23.6 (21.2-26.2)	20.9	19.6
<u>#3.1 宗教を信じるか</u>				支持政党なし	59.7 (56.7-62.6)	62.3	63.8
信じている	30.3 (27.1-33.6)	27.1	26.6	<u>#9.6 日本人・西洋人の優劣</u>			
信じていない	69.7 (66.4-72.9)	72.9	73.4	すぐれている	31.2 (28.2-34.4)	28.9	28.1
<u>#3.2b 「宗教心」は大切か</u>				劣っている	6.8 (5.4-8.5)	7.9	8.3
大切	70.1 (67.3-72.7)	66.7	65.8	同じだ	30.6 (27.7-33.7)	31.2	30.8
大切でない	14.8 (12.9-17.0)	16.9	17.2	ひとくちではいいない	24.4 (21.8-27.3)	24.9	25.3
<u>#4.1 先祖を尊ぶか</u>				(参考)			
尊ぶ方	59.1 (55.9-62.2)	55.5	54.8	<u>#2.12c 人は信頼できるか</u>			
普通	31.5 (28.4-34.7)	33.6	34.3	信頼できる	32.6 (29.6-35.7)	31.7	30.9
尊ばない方	8.9 (7.2-10.9)	10.2	10.2	用心する	59.2 (55.9-62.5)	60.2	61.1
<u>#4.32 離婚すべきでないか</u>				<u>#2.12 他人のためか自分のためか</u>			
離婚すべきでない	27.2 (24.4-30.2)	24.3	24.6	他人の役に	34.0 (31.1-37.0)	33.8	33.0
ひどい場合には離婚	43.6 (40.5-46.8)	44.3	43.6	自分のことだけ	50.8 (47.6-53.9)	50.8	51.3
合意さえあれば離婚	26.9 (24.4-29.7)	29.2	29.3				

表6 人口統計学的属性変数の分布

項目 カテゴリ	母集団	計画	K型		M型	
			回収	返送	回収	返送
合計(人数)	95,839,000	4,193	1,192	579	1,158	590
性						
男性	49.2	49.9	46.2	44.4	46.6	46.4
女性	50.8	50.1	53.8	55.6	53.4	53.6
年齢層						
20歳代	17.3	17.3	10.2	7.4	12.2	7.1
30歳代	18.6	17.8	17.5	13.8	14.8	13.7
40歳代	16.2	16.4	16.4	17.8	15.8	15.1
50歳代	19.9	20.4	22.5	21.8	21.3	24.1
60歳代	16.3	17.0	21.2	24.7	20.5	24.7
70歳代	11.7	10.8	12.2	14.5	15.5	15.3
区市郡						
区部	***	15.6	12.4	13.6	13.0	12.4
市(50万~)	***	13.0	12.3	13.6	12.2	12.4
市(20万~)	***	19.2	18.7	18.7	18.2	20.0
市(10万~)	***	13.0	13.5	13.1	12.3	12.7
市(5万~)	***	13.3	13.8	13.6	14.2	13.9
市(~5万)	***	5.4	6.5	6.0	6.2	4.9
郡部	***	20.4	22.7	21.2	23.9	23.7
地方						
北海道	4.5	4.7	4.8	4.7	5.7	5.4
東北	7.5	7.8	8.1	8.3	8.8	9.8
関東	32.8	32.3	29.4	30.4	28.8	28.3
中部(東)	7.2	7.3	8.6	10.2	7.7	7.8
中部(西)	9.7	9.6	10.4	10.0	10.6	11.0
近畿	17.9	17.6	15.4	15.0	15.5	17.3
中国	5.9	6.3	7.2	7.4	6.4	6.8
四国	3.2	3.0	3.4	3.8	3.2	3.9
九州	11.2	11.5	12.9	10.2	13.4	9.7
世帯人数						
1人	***	***	8.4	7.1	7.3	6.1
2人	***	***	23.2	26.9	24.5	27.5
3人	***	***	20.1	18.5	22.5	23.4
4人	***	***	23.9	22.8	22.1	22.4
5人	***	***	12.1	11.9	11.6	9.3
6人以上	***	***	12.3	12.8	12.0	11.4
住居形態						
一戸建	***	70.8	77.5	80.5	80.3	82.7
その他	***	29.2	22.5	19.5	19.7	17.3
学歴						
小学	}22.1	***	3.2	2.4	4.7	2.9
中学		***	19.1	18.7	20.2	20.2
高校	49.4	***	48.2	48.5	47.5	48.1
大学	28.5	***	29.5	30.4	27.6	28.8
就業状態						
就業	64.2	***	64.4	60.6	64.0	63.6
その他	35.8	***	35.6	39.4	36.0	36.4

注) 母集団の学歴と就業状態は平成12年国勢調査による。

他は、平成15年10月1日現在の推計人口(総務省統計局)による。