

調査への指向性変数を用いた調査不能 バイアスの二段補正

—「日本人の国民性 第12次全国調査」への適用—

土屋 隆裕[†]

(受付 2009年9月30日; 改訂 12月21日; 採択 12月24日)

要 旨

「日本人の国民性 第12次全国調査」の回収率は51.6%となり、過去最低を記録した。国民性調査では各回答選択肢の標本割合を母集団推定値としている。しかし回収率の低さゆえ、項目によっては調査不能によるバイアスは小さくないと考えられる。本稿ではウェイトの調整を行うことで、調査不能バイアスの軽減を試みた。調査不能の二大理由は不在と拒否である。そこでウェイトのキャリブレーションを各不能理由に応じた二段階で行うこととした。このときウェイト調整の補助変数としては、人口統計学的な属性変数だけでなく、調査に対する態度を表す変数(調査への指向性変数と呼ぶ)も用いることを提案した。ウェイトを調整した推定値は、属性変数だけを用いて調整を行った場合よりも、指向性変数も用いて調整を行った場合の方が、標本割合との差が大きく、10ポイント近い差を見せるカテゴリもあった。母集団における真の割合が未知である以上、本稿で提案したウェイト調整法によって不能バイアスが消滅したとは断じ得ない。しかしこれまでに徐々に明らかになってきた調査不能者の特徴に鑑みて、得られた推定値は少なくとも標本割合に比べればより妥当なものと考えられた。

キーワード：調査不能、不在、拒否、キャリブレーション、調査への指向性変数。

1. はじめに

1.1 回収率の低下と調査不能バイアス

2008年秋に実施された「日本人の国民性 第12次全国調査」(中村 他, 2009)では、抽出標本の6,398に対し回収標本は3,302となり、回収率は51.6%となった。この値は、それまで過去最低であった前回第11次調査(坂元 他, 2004)の56.0%をさらに下回り、調査環境の悪化が依然として続く状況を改めて顕わにした。

日本人の国民性調査では、原則として全ての個人が等しい抽出ウェイト(土屋, 2009, p. 33)を持つ自己加重標本が抽出されるよう標本設計を行っている。そのため国民性調査の結果は伝統的に、回収標本における割合(以下、標本割合)をそのまま公表してきた。しかし国民性調査の目標母集団は日本人成人全体(ただし、最近2回は20歳以上80歳未満)である。抽出標本の半数近くが調査不能であれば、標本割合は母集団割合の推定値としてバイアスを持つ場合が少なくないと考えられる(土屋, 2006)。本稿では推定のためのウェイトの調整(Oh and Scheuren,

[†] 統計数理研究所：〒190-8562 東京都立川市緑町 10-3

表 1. 調査不能理由の分布 (第 12 次調査).

死亡 10 (0.3%)	移転 288 (9.3%)	該当者なし 70 (2.3%)	尋ね当たらず 48 (1.6%)	長期不在 101 (3.3%)	
病気 138 (4.5%)	一時不在 892 (28.8%)	拒否 1,501 (48.5%)	老衰 21 (0.7%)	その他 27 (0.9%)	合計 3,096

1983) を行い, 調査不能によるバイアスの軽減を試みる.

1.2 本稿の目的

近年の国民性調査において推定用のウェイトを調整し, 調査不能バイアスの軽減を試みることは, 既に土屋 (2005) や土屋 (2006) で行われている. 特に土屋 (2006) は, 少数の人口統計学的属性変数だけを用いてウェイト調整を行うと推定値の誤差はかえって大きくなるおそれがあること, 誤差を縮小するにはなるべく多くの補助変数を同時に用いるべきであること, ウェイトの調整方法の選択よりも調整に用いる補助変数の内容の方が重要であることを明らかにした. そしてウェイト調整の補助変数の一つとして調査への指向性変数を用いることを提案した. ここで言う調査への指向性変数とは, 調査に対する回答者のいわば態度等を表す変数であり, 回答者の外形的な属性とは異なる変数のことである. その具体的な内容や考え方は次節で詳述する.

ところで土屋 (2006) では調査不能理由として拒否のみを考慮し, 拒否に対応した指向性変数一つだけを用いてウェイト調整を行っている. しかし現実の調査不能理由としては, 拒否に加えて不在も少なくない. 例えば表 1 は第 12 次調査の調査不能理由の内訳である. 拒否が 48.5% と不能の半分を占める. 同時に移転は 9.3% であり, 長期不在と一時不在を合わせた不在が 32.1% というように, 調査不能の 4 割はそもそも調査員が対象者に会うことすらできていない.

ウェイトの調整に当たっては, このような拒否と不在という調査不能の二大理由のそれぞれに対応した指向性変数を用いる必要がある. 本稿では複数の指向性変数を用いたウェイト調整を試み, 標本割合や, 属性変数のみを用いたウェイト調整の結果と比較することで, その妥当性を検討する.

2. ウェイト調整の方法

2.1 ウェイトのキャリブレーション

本稿で用いるウェイト調整の方法は以下のとおりである. この方法は, 事後層化やレイキング, 一般化回帰推定量をその特別な場合として含む非常に一般的な方法である (Deville and Särndal, 1992; Deville et al., 1993; 土屋, 2009).

母集団 pop における総計 τ_j が知られている補助変数, あるいはその不偏推定値 $\hat{\tau}_j$ が得られる補助変数が合わせて J 個あるものとする. このとき第 i 回答者の調整前のウェイトを w_i とすると, 調整後のウェイト w_i^c は各補助変数に関して以下が成り立つようにすべきである.

$$(2.1) \quad \hat{\tau}_j^c = \sum_{i \in \text{res}} w_i^c x_{ij} = \begin{cases} \sum_{i \in \text{pop}} x_{ij} = \tau_j \\ \text{あるいは} \\ \sum_{i \in \text{samp}} w_i x_{ij} = \hat{\tau}_j \end{cases}, \quad j = 1, \dots, J$$

ただし抽出標本 samp のうちの回収標本を res とし, 第 i 回答者の第 j 補助変数の値を x_{ij} と

する．本稿ではカテゴリカルな補助変数のみを用いるので，第 j 補助変数とは第 j カテゴリのことであり， x_{ij} は第 i 回答者が第 j カテゴリに該当すれば 1，そうでなければ 0 という値をとる二値変数である．

(2.1) 式が要求していることは，母集団総計 τ_j あるいはその不偏推定値 $\hat{\tau}_j$ が分かる補助変数に関しては，調整後のウェイト w_i^c を用いた推定値 $\hat{\tau}_j^c = \sum_{i \in \text{res}} w_i^c x_{ij}$ がそれらの値に一致すべきということである．例えば性別や年齢層，地域などは国勢調査等の結果から目標母集団における真の総数 τ_j が知られている．ウェイト調整後の推定値 $\hat{\tau}_j^c$ はそれら真値 τ_j に一致すべきである．また住居形態などは，目標母集団における真値 τ_j が分からなくとも，不偏推定値 $\hat{\tau}_j = \sum_{i \in \text{samp}} w_i x_{ij}$ は求まる．調査員が各対象者の住居形態を記録することによって，抽出標本 samp 全体から変数値 x_{ij} が得られるからである．調整後のウェイト w_i^c を用いて回収標本 res だけから求めた推定値 $\hat{\tau}_j^c$ は，不偏推定値 $\hat{\tau}_j$ に一致すべきである．なお Särndal and Lundström (2005) は真値 τ_j が分かる場合を infoU と呼び，不偏推定値 $\hat{\tau}_j$ を用いる場合を infoS と呼んでいる．一般に我が国における社会調査では，infoU, infoS とともに入手可能な補助変数の数は限られている．利用できる補助変数は選択の余地がないことが多く，可能な限り多数の補助変数を用いることになる．

さらに調整後のウェイト w_i^c は，調整前のウェイト w_i と大きくは異ならない方が望ましい．そこで (2.1) 式が成り立つという条件の下で， w_i と w_i^c の間の距離関数 $G(w_i, w_i^c)$ を用いた $\sum_{i \in \text{res}} w_i G(w_i, w_i^c)$ が最小となる w_i^c を求める．本稿では $G(w_i, w_i^c)$ として以下を用いる (Folsom, 1991; Folsom and Singh, 2000)．

$$(2.2) \quad G(w_i, w_i^c) = \frac{1}{A} \left\{ \left(\frac{w_i^c}{w_i} - L \right) \log \frac{w_i^c/w_i - L}{C - L} + \left(U - \frac{w_i^c}{w_i} \right) \log \frac{U - w_i^c/w_i}{U - C} \right\}$$

ただし U と L は w_i^c/w_i の上限と下限であり，あらかじめ定めておく値である． C も $L < C < U$ を満たす定数である．また $A = (U - L) / \{(U - C)(C - L)\}$ である．距離関数としてはより簡単な線形関数 $G(w_i, w_i^c) = (w_i^c/w_i - 1)^2/2$ もあるが，調整後のウェイト w_i^c が負の値となることがある．(2.2) 式では， $L > 0$ としておけば負の w_i^c が得られることはない．そのため (2.2) 式を用いる． w_i^c を求めるにはニュートン-ラフソン法を用いればよい (Folsom and Singh, 2000)．

求められた w_i^c は，あるベクトル β を用いて次式のように表せる．

$$(2.3) \quad w_i^c = w_i g_i = w_i \frac{L(U - C) + U(C - L) \exp(Ax_i' \beta)}{(U - C) + (C - L) \exp(Ax_i' \beta)}$$

ただし x_i は x_{i1}, \dots, x_{iJ} を並べたベクトルである．(2.3) 式の g_i は g -ウェイトと呼ばれる．なお $L=0, C=1, U \rightarrow \infty$ とすれば，(2.3) 式による w_i^c はレイキングによるウェイトに一致し， $L=1, C=2, U \rightarrow \infty$ とすれば， w_i^c は次式となることが知られている．

$$(2.4) \quad w_i^c = w_i \frac{1 + \exp(-x_i' \beta)}{\exp(-x_i' \beta)}$$

本稿では $L=0, C=1, U \rightarrow \infty$ としたので，レイキングを行ったことになる．なお，これらの値は変えてみても結果に大きな違いは見られなかった．土屋 (2006) も距離関数を変えてキャリブレーションを行ったが，結果に違いはほとんど見られなかったとしている．

調整後のウェイト w_i^c を用いると，質問項目の第 k カテゴリの割合の推定値は $\hat{p}_k = 100 \times \sum_{i \in \text{res}} w_i^c y_{ik} / \sum_{i \in \text{res}} w_i^c$ となる．ただし y_{ik} は第 i 回答者が第 k カテゴリに該当すれば 1，そうでなければ 0 という値をとる二値変数である．

2.2 二段階のキャリブレーション

一般に近年の調査不能の二大理由は拒否と不在である．第 12 次調査ではこの 2 つで調査不能の 8 割を占める．したがって不能バイアスの補正に当たっては，拒否と不在の 2 つを軸に考

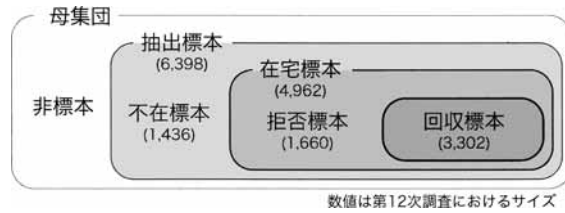


図 1. 母集団・抽出標本・在宅標本・回収標本の間関係.

えていけばよい (Groves and Couper, 1998). 対象者は、まず調査員が訪問したとき不在か在宅かで 2 群に分かれ、在宅であれば次に拒否か回収かで 2 群に分かれると考えるのである. 図 1 には標本における不在と拒否、回収の間関係を示した.

本稿では表 1 に示す 10 の調査不能理由を次のように大きく 2 つに分類する. すなわち病氣, 拒否, 老衰を広い意味での拒否とし, 残りの 7 つを広い意味での不在とする. 前者はいずれも対象者本人が在宅していた可能性が高い理由だからである. これに対し後者は全て, 調査員の訪問時には対象者本人が不在だった理由である. この基準に従って第 12 次調査の標本を分類すると図 1 のとおりとなる. 抽出標本 6,398 のうち広義の不在標本は 1,436 であり, 残りの 4,962 を在宅標本と称することとする. 在宅標本 4,962 のうち回収標本は 3,302 であり, 残りの 1,660 は広義の拒否標本である.

ウェイトのキャリブレーションは二段階で行う. 具体的には以下のとおりである. まず広義の拒否による調査不能を補正する. すなわち標本抽出デザインを反映した抽出ウェイト

$$(2.5) \quad w_i = \frac{N}{n} = \begin{cases} 96,669,010/3,211 & (\text{K 型調査票}) \\ 96,669,010/3,187 & (\text{M 型調査票}) \end{cases}$$

を (2.2) 式あるいは (2.3) 式における w_i とする. N は母集団サイズ, n は抽出標本サイズである. そして g-ウェイト $g_i^{\text{拒}}$ すなわちキャリブレーションウェイト $w_i^c = w_i g_i^{\text{拒}}$ を求める. このとき用いる τ_j や $\hat{\tau}_j$ は母集団のうち調査をしたとき不在ではない人々から成る部分母集団における真値, あるいはその不偏推定値である. w_i^c を用いた推定値 $100 \times \sum_{i \in \text{res}} w_i^c y_{ik} / \sum_{i \in \text{res}} w_i^c$ は, この部分母集団における第 k カテゴリの割合の推定値となる.

次に広義の不在による調査不能を補正する. すなわち上記で求めたキャリブレーションウェイト w_i^c を改めて (2.2) 式あるいは (2.3) 式における w_i として用いて, 新たに g-ウェイト $g_i^{\text{不}}$ を求めるのである. このときの τ_j や $\hat{\tau}_j$ は目標母集団全体における真値あるいはその不偏推定値である. (2.5) 式の w_i を用いると, 最終的なウェイトは $w_i^{\text{不}} = w_i^c g_i^{\text{不}} = w_i g_i^{\text{拒}} g_i^{\text{不}}$ となる.

ところで広義の拒否の補正に用いる補助変数は, 調査をしたとき不在とはならない人々から成る部分母集団に関して真値 τ_j が分かっている変数, あるいはその不偏推定値 $\hat{\tau}_j$ が得られる変数でなければならない. 真値 τ_j が分かる補助変数は入手することができない. 「不在とはならない人々」に関して国勢調査などの結果は公表されていないからである. しかし表 2 に示す 5 つの変数に関しては, 不偏推定値 $\hat{\tau}_j$ が得られる. 在宅標本を含む抽出標本全体について変数値 x_{ij} が得られているからである. つまり在宅標本を home とすると, (2.1) 式における $\hat{\tau}_j$ を

$$(2.6) \quad \hat{\tau}_j = \sum_{i \in \text{home}} w_i x_{ij}, \quad j = 1, \dots, J$$

とする. また (2.6) 式の補助変数 x としては, 表 2 の 5 変数全てを用いることとする. 土屋 (2006) によれば, なるべく多数の補助変数を用いるべきだからであり, 在宅標本あるいは抽出

表 2. 拒否と不在のいずれの補正においても用いた属性変数.

	抽出標本	在宅標本	回収標本
<u>全体</u>	6,398 (100%)	4,962 (100%)	3,302 (100%)
<u>性別</u>			
男性	3,193 (49.9%)	2,376 (47.9%)	1,562 (47.3%)
女性	3,205 (50.1%)	2,586 (52.1%)	1,740 (52.6%)
<u>年齢層</u>			
20 歳代	911 (14.2%)	522 (10.5%)	353 (10.7%)
30 歳代	1,189 (18.6%)	808 (16.3%)	533 (16.1%)
40 歳代	1,091 (17.1%)	849 (17.1%)	540 (16.4%)
50 歳代	1,335 (20.9%)	1,104 (22.2%)	729 (22.1%)
60 歳代	1,156 (18.1%)	1,020 (20.6%)	700 (21.2%)
70 歳代	716 (11.2%)	659 (13.3%)	447 (13.5%)
<u>市郡</u>			
6 大都市	1,024 (16.0%)	714 (14.4%)	390 (11.8%)
市部・人口 50 万以上	896 (14.0%)	678 (13.7%)	440 (13.3%)
市部・人口 20 万以上	1,371 (21.4%)	1,081 (21.8%)	724 (21.9%)
市部・人口 10 万以上	970 (15.2%)	752 (15.2%)	481 (14.6%)
市部・人口 5 万以上	855 (13.4%)	676 (13.6%)	465 (14.1%)
市部・人口 5 万未満	433 (6.8%)	348 (7.0%)	258 (7.8%)
町村	851 (13.3%)	713 (14.4%)	544 (16.5%)
<u>地方</u>			
北海道	288 (4.5%)	240 (4.8%)	160 (4.8%)
東北	450 (7.0%)	375 (7.6%)	279 (8.4%)
関東	2,162 (33.8%)	1,542 (31.1%)	957 (29.0%)
中部(東)	480 (7.5%)	385 (7.8%)	297 (9.0%)
中部(西)	576 (9.0%)	481 (9.7%)	337 (10.2%)
近畿	1,151 (18.0%)	886 (17.9%)	535 (16.2%)
中国	399 (6.2%)	311 (6.3%)	220 (6.7%)
四国	193 (3.0%)	165 (3.3%)	112 (3.4%)
九州	699 (10.9%)	577 (11.6%)	405 (12.3%)
<u>住居形態</u>			
一戸建	4,377 (68.4%)	3,644 (73.4%)	2,543 (77.0%)
集合住宅	1,351 (21.2%)	913 (18.4%)	546 (16.5%)
オートロック集合住宅	502 (7.8%)	323 (6.5%)	180 (5.5%)
その他・D.K.	168 (2.6%)	82 (1.7%)	33 (1.0%)

標本全体について値が得られている人口統計学的属性変数はこの 5 つに限られるからである。 J は 5 変数のカテゴリ総数 $J=28$ となる。

また広義の不在の補正に用いる補助変数は、目標母集団全体における真値 τ_j あるいはその不偏推定値 $\hat{\tau}_j$ が得られる変数である必要がある。そこで、そのような条件に合致し入手可能な全ての変数として、表 2 と表 3 の 10 変数の合計 $J=59$ カテゴリを用いる。ただし表 3 の母集団割合のうち、学歴は平成 12 年国勢調査による値であり、残りは平成 17 年国勢調査による。さらに婚姻は 17 歳以上 76 歳以下日本人人口によるが、勤務形態は 20 歳以上 79 歳以下、残りは全ての年齢における割合である。

なお Iannacchione (2003) は傾向スコア (Rosenbaum and Rubin, 1983; 星野・繁榊, 2004) を用いてまず不在を補正したウェイトを求め、次にこのウェイトに対して拒否を補正したウェイトをさらに求めている。本稿では最初に拒否を補正した後不在を補正する。これは次節で述べ

表 3. 不在の補正にのみ用いた属性変数.

	母集団	回収標本
<u>全体</u>	96,669,010	3,302 (100%)
<u>婚姻</u>		
未婚	27.9%	488 (14.8%)
死別	4.6%	186 (5.6%)
離別	4.6%	143 (4.3%)
既婚	62.8%	2,485 (75.3%)
<u>学歴</u>		
小学校・中学校	25.0%	572 (17.3%)
高等学校	47.2%	1,564 (47.4%)
短大・大学・大学院	27.9%	1,166 (35.3%)
<u>勤務形態</u>		
主に仕事をした	53.5%	1,864 (56.5%)
家事などのかたわら仕事をした	8.9%	340 (10.3%)
通学のかたわら仕事をした	0.7%	27 (0.8%)
仕事を休んでいた	1.0%	65 (2.0%)
仕事を探していた	4.0%	34 (1.0%)
家事をした	17.9%	593 (18.0%)
通学	1.8%	26 (0.8%)
その他	12.2%	353 (10.7%)
<u>住居の種類</u>		
持ち家	70.7%	2,650 (80.3%)
都道府県・市区町村営の賃貸住宅	4.1%	130 (3.9%)
都市再生機構・公社等の賃貸住宅	1.8%	47 (1.4%)
民営の賃貸住宅	18.9%	373 (11.3%)
給与住宅	2.8%	53 (1.6%)
住宅に間借り	0.9%	20 (0.6%)
その他	0.8%	29 (0.9%)
<u>世帯人員数</u>		
1人	11.6%	249 (7.5%)
2人	20.8%	845 (25.6%)
3人	22.1%	783 (23.7%)
4人	24.7%	716 (21.7%)
5人	11.4%	397 (12.0%)
6人	5.8%	172 (5.2%)
7人	2.6%	106 (3.2%)
8人	0.8%	25 (0.8%)
9人以上	0.3%	9 (0.3%)

注 1) 学歴は平成 12 年国勢調査により、他の属性は平成 17 年国勢調査による。また婚姻は 17 歳以上 76 歳以下日本人人口により、勤務形態は 20 歳以上 79 歳以下、他は全ての年齢における割合である。

注 2) 婚姻の‘その他’と‘D.K.’は‘既婚’とした。

注 3) 学歴の‘その他’は‘小学校・中学校’とした。

注 4) 世帯人員数の‘その他’は‘1人’とした。

るように、調査をしたとき不在とはならない人々から成る部分母集団に対して、指向性変数を用いたキャリブレーションを行うためである。

2.3 調査への指向性変数

ウェイトの調整に属性変数だけを用いると、推定値のバイアスはかえって大きくなるおそれ

表 4. 拒否の補正に用いた指向性変数.

#1.90* 調査協力理由 [リスト] 最後に、今回調査にご協力いただいた理由はどういったことでしょうか。あなたの気持ちに一番近いものをお選び下さい？	
1. 役に立ちたいと思ったから	540 (16.4%)
2. 内容が面白そうだったから	89 (2.7%)
3. 公的な研究機関が行う調査だから	782 (23.7%)
4. 今までに、こういう調査に協力したことがほとんどなかったから	324 (9.8%)
5. 特に断る理由もなかったから	1,434 (43.4%)
6. その他 [記入]	88 (2.7%)
7. D.K.	45 (1.4%)
#1.91 再協力の意向 [リスト] 仮にこういった調査がもう一度あるとしたら、あなたはどのようにしますか？	
1. 必ず答えるようにしたい	645 (19.5%)
2. なるべく答えるようにしたい	1,596 (48.3%)
3. あまり答えたくない	700 (21.2%)
4. 答えたくない	270 (8.2%)
5. その他 [記入]	27 (0.8%)
6. D.K.	64 (1.9%)
#1.92a 対象者の反応 (調査員記入) 調査に対する対象者の反応はどうでしたか。	
1. 最初から好意的	2,017 (61.1%)
2. だんだん好意的になった	942 (28.5%)
3. だんだん迷惑そうになった	163 (4.9%)
4. 最初から迷惑そう	180 (5.5%)

表 5. 不在の補正に用いた指向性変数.

#1.80 休日は在宅か [リスト] 休日は、あなたはご在宅のことが多いですか、それともお出かけのことが多いですか？	
1. ほとんど在宅している	766 (23.2%)
2. どちらかといえば在宅しているほうが多い	886 (26.8%)
3. どちらかといえば出かけているほうが多い	1,276 (38.6%)
4. ほとんど家にはいない	288 (8.7%)
5. その他 [記入]	46 (1.4%)
6. D.K.	42 (1.3%)

がある。例えば若年層の男性は一般に不在がちである。家にいることが多いゆえに回収できた若年男性に大きなウェイトを与えると、項目によってはかえって誤差が大きくなってしまふのである。そこでウェイト調整のためには、調査をしたときにどの程度不在となりにくい、あるいは積極的な協力の得られやすさを示す変数を補助変数として追加する必要がある。本稿ではそのような変数を調査への指向性変数と呼ぶ。表 4 には拒否の補正時に用いる指向性変数、表 5 と表 6 には不在の補正時に用いる指向性変数を示す。

ウェイト調整のために指向性変数を用いるには、指向性変数に関しても (2.1) 式の τ_j あるいは $\hat{\tau}_j$ が必要である。表 4 や表 5 の項目に対する回答は、回収標本についてのみ得られており、拒否標本や不在標本については得られていない。しかし例えば表 4 の“#1.91 再協力の意向”について考えてみよう。この項目に対する回答は拒否標本からは得られていないが、仮に得ら

表 6. 訪問回数の分布の比較.

	抽出標本	在宅標本	回収標本
0回	106 (1.7%)	86 (1.7%)	0 (0.0%)
1回	1,797 (28.1%)	1,516 (30.6%)	983 (29.8%)
2回	1,594 (24.9%)	1,448 (29.2%)	1,076 (32.6%)
3回	1,184 (18.5%)	964 (19.4%)	664 (20.1%)
4回	737 (11.5%)	507 (10.2%)	324 (9.8%)
5回	513 (8.0%)	261 (5.3%)	169 (5.1%)
6回以上	467 (7.3%)	180 (3.6%)	86 (2.6%)

れたとしても‘必ず答えるようにしたい’あるいは‘なるべく答えるようにしたい’とはならないであろう。拒否標本の人々は、‘答えたくない’あるいはせいぜい‘あまり答えたくない’と回答するはずである。‘#1.92a 対象者の反応(調査員記入)’についても、拒否標本は‘最初から好意的’なはずはない。つまり表4の変数に関して拒否標本からの回答は得られていないが、合理的な推論によって拒否標本の回答の補定は可能である。

そこで表4の3つの項目はそれぞれ4つから7つのカテゴリを持つが、‘#1.90* 調査協力理由’と‘#1.91 再協力の意向’についてはカテゴリ1と2の二つをまとめて一つのカテゴリとし、残りのカテゴリをまとめてもう一つのカテゴリとする。また‘#1.92a 対象者の反応(調査員記入)’については‘最初から好意的’とそれ以外の2カテゴリとする。そして拒否標本はいずれの項目についても後者のカテゴリに該当するものとする。例えば‘#1.91 再協力の意向’であれば、在宅標本は全体で4,962であるが、そのうち回収標本の645 + 1,596 = 2,241は再協力の意向有りの方のカテゴリに該当するものとし、回収標本の残り700 + 270 + 27 + 64 = 1,061と拒否標本1,660の合わせて1,061 + 1,660 = 2,721は再協力の意向無しの方のカテゴリに該当するものとする。抽出標本は自己加重標本なので、在宅標本における割合2,241/4,962 = 45.2%と2,721/4,962 = 54.8%が、調査をしたとき不在とはならない人々から成る部分母集団における不偏推定値となる。これらの割合に部分母集団サイズの推定値 $4,962w_i$ を乗じた値を(2.1)式における $\hat{\tau}_j$ として用いればよい。ただし w_i は(2.5)式の抽出ウェイトである。

Stinchcombe et al. (1981)やHolt and Elliot (1991), Proctor (1996)は、回収標本のうち例えば容易に協力が得られなかった者など、調査不能者に似ていると考えられる回答者の質問項目への回答を、調査不能者の回答として補定することを提案している。しかし協力が得難かったからといって、質問に対する回答までもが必ずしも調査不能者と同じであるとは限らない。指向性変数の考え方は、目的とする質問項目の回答を直接補定するのではなく、不在であったり調査に対して拒否したりするという調査不能者の特徴や行為そのものを補定値とすることができる変数を用意し、補定した変数を補助変数としてウェイト調整に用いようとするものである。

表5に示す‘#1.80 休日は在宅か’は、回収標本における標本割合である。表6に示す訪問回数を見ると、抽出標本全体と在宅標本とでは抽出標本の方が訪問回数がやや多いのに対し、在宅標本と回収標本とでは訪問回数にほとんど違いがない。そのため表5の分布を在宅標本における分布として用いても問題は少ないであろう。‘ほとんど在宅している’と‘どちらかといえば在宅しているほうが多い’の2カテゴリをまとめ、残りの4つのカテゴリもまとめた上で、不在標本は後者のカテゴリに該当するものとする。母集団において在宅がちのカテゴリに該当する人の割合は $(766 + 886)/3,302 \times 4,962/6,398 = 38.8\%$ とし、不在がちのカテゴリに該当する人の割合は $100 - 38.8 = 61.2\%$ とする。これらに母集団サイズ96,669,010を乗じた値が $\hat{\tau}_j$ である。表6の訪問回数の母集団における分布は抽出標本における分布をそのまま用いる。ただし‘0回’は‘1回’にまとめた。

3. 結果

3.1 全体的な比較

表 2 と表 3 の属性変数に加えて、表 4 から表 6 の指向性変数を用いて二段階のキャリブレーションを行ったときの推定値を \hat{p}_k^c とする。なお本稿の目的は指向性変数の効果を確かめることなので、表 2 と表 3 の属性変数だけを用いて、指向性変数を用いずに二段階のキャリブレーションを行ったときの第 k カテゴリの割合の推定値 \hat{p}_k^{c*} も同時に求めておく。

まず図 2 は、属性変数のみによる推定値 \hat{p}_k^{c*} と標本割合 \hat{p}_k との差の絶対値 $|\hat{p}_k^{c*} - \hat{p}_k|$ のうち項目ごとの最大値を横軸とし、指向性変数を加えた推定値 \hat{p}_k^c と標本割合 \hat{p}_k との差の絶対値 $|\hat{p}_k^c - \hat{p}_k|$ のうち項目ごとの最大値を縦軸として、基本的に全ての項目をプロットしたものである。例えば表 7 にある“#2.3c 家庭に満足か”について言えば、‘満足’から‘D.K.’まで 6 つのカテゴリのうち標本割合と属性変数だけによる推定値との差が最大である‘満足’の $|38.5 - 41.6| = 3.1$ を横軸とし、標本割合と指向性変数を用いた推定値との差が最大である同じく‘満足’の $|36.5 - 41.6| = 5.1$ を縦軸として、この項目をプロットした。ただし“#1.1 性別”や“#1.2 年齢”など、# 番号が #1. で始まる基本項目についてはプロットしていない。結局、プロットされたのは K 型調査票では 66 項目、M 型調査票では 57 項目である。

図 2 によれば、明らかに指向性変数を追加してキャリブレーションを行う方が推定値の変化が大きい。属性変数のみでは、標本割合と比べてほとんどの項目がたかだか 4 ポイント程度しか推定値は変わらないのに対し、指向性変数を追加すると 6 ポイント近くまで変わり、10 ポイント近く変わる項目もある。

ところで、土屋(2006)は“#1.90* 調査協力理由”のみを指向性変数として一段階でウェイトのキャリブレーションを行っているのに対し、本稿では指向性変数を増やし二段階でウェイトのキャリブレーションを行うことを提案している。両者の間で、推定値に与える影響の大きさを比較したものが図 3 である。まず図 3 の縦軸の値は、項目ごとに図 2 の縦軸の値 $|\hat{p}_k^c - \hat{p}_k|$ から横軸の値 $|\hat{p}_k^{c*} - \hat{p}_k|$ を引いたものである。先に例示した“#2.3c 家庭に満足か”であれば、 $|36.5 - 41.6| - |38.5 - 41.6| = 5.1 - 3.1 = 2.0$ が図 3 の縦軸の値である。つまり図 2 における各点が対角線から上にどの程度ズレているかを示したものである。図 3 の縦軸ではほとんどの項目が正の値となっており、属性変数だけを用いたときの推定値の変化に比べ、指向性変数を二段階で追加したときの変化の方が大きいことが改めて分かる。次に図 3 の横軸の値は縦軸と同様に求めたものであるが、二段階でキャリブレーションを行った推定値 \hat{p}_k^c の代わりに、“#1.90*

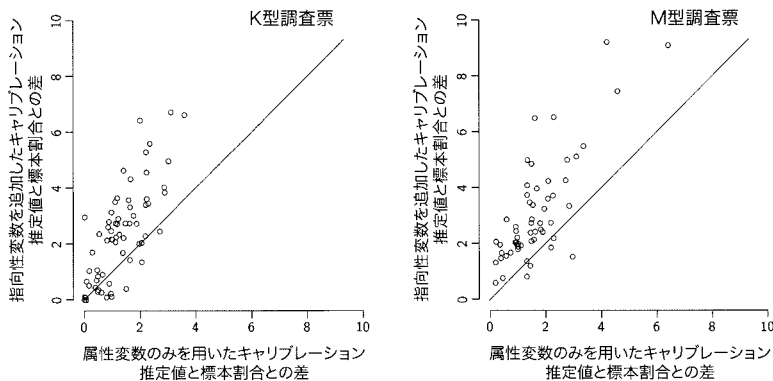


図 2. 属性変数のみを用いた推定値と指向性変数を追加した推定値の比較。

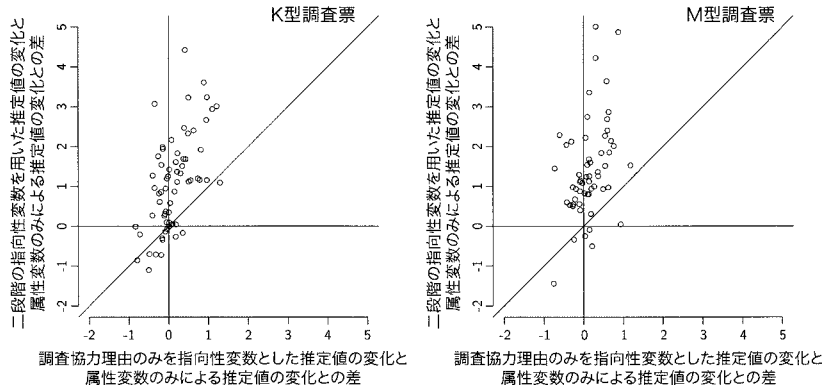


図3. 調査協力理由のみを指向性変数とした推定値の変化と二段階の指向性変数を用いた推定値の変化との比較.

調査協力理由”と表2, 表3の全部で11変数を用いて一段階でキャリブレーションを行った推定値を用いた. 例えば“#2.3c 家庭に満足か”であれば, ‘満足’の推定値は37.8%であったので, $|37.8 - 41.6| - |38.5 - 41.6| = 3.8 - 3.1 = 0.7$ が横軸の値である. 図3の特にM型調査票では, 横軸で多くの項目が正の値となっている. 指向性変数として追加する変数が“#1.90* 調査協力理由”の一つだけであっても, 標本割合との差は属性変数だけの場合よりもさらに1から2ポイント程度広がるのが分かる. しかし図3ではほとんどの項目が対角線より上にある. 土屋(2006)の方法に比べ, 本稿で提案した二段階のキャリブレーションの方が推定値をより大きく変えることになる.

3.2 項目ごとの比較

次に表7は, 指向性変数を二段階で加えた推定値 \hat{p}_k^c と標本割合 \hat{p}_k との差の絶対値 $|\hat{p}_k^c - \hat{p}_k|$ が5ポイント以上あるカテゴリを含む項目について, 標本割合 \hat{p}_k , 属性変数だけを用いた二段階のキャリブレーション推定値 \hat{p}_k^{c*} , 指向性変数を加えた二段階のキャリブレーション推定値 \hat{p}_k^c を並べたものである. 属性変数だけを用いた推定値 \hat{p}_k^{c*} が標本割合 \hat{p}_k よりも小さいカテゴリのほとんどは, 指向性変数を加えることで推定値 \hat{p}_k^c がより小さくなっており, 逆に $\hat{p}_k^c > \hat{p}_k$ であるカテゴリのほとんどは $\hat{p}_k^c > \hat{p}_k^{c*} > \hat{p}_k$ となっている.

個々の項目について見ていくと, 標本割合に比べて指向性変数を追加した推定値(以下, 補正值)は, ‘世の中のためになる’ことをしたい(#2.10)や‘人のためになること’をしたい(#2.11)という割合が減り, 逆に自分が‘しあわせにくらす’ことが第一(#2.10)や‘自分の好きなこと’をしたい(#2.11)という割合が増える. 特に“#2.11 好きなくらし方か人のためか”については, 標本割合では‘人のためになること’(54.1%)の方が‘自分の好きなこと’(39.6%)よりも15ポイント近く多いのに対し, 補正值では両者は拮抗する. また“#2.35 ボランティア活動”では‘今後もしない’の標本割合27.7%が補正值では36.8%と上昇し, “#9.16 国際貢献”では自分たちの‘生活水準上げる’の標本割合が55.9%から61.4%へと上昇する.

他人はスキがあればあなたを‘利用しようとしている’(#2.12b)や, たいていの人は‘用心する’にこしたことはない(#2.12c)という回答は, 標本割合よりも補正值の方が大きい. また‘宗教否定’(#3.6)や‘両方否定’(#3.6)は補正值の方が大きく, 「宗教心」は‘大切でない’(#3.2b)も補正值の方が大きい. ‘給料が多い会社’につとめたい(#5.6b)という割合やお金があれば‘仕事がなくともよい’(#7.25)という割合も補正值の方が大きい. さらに補正值では,

表 7. 3 種類の推定値の比較.

項目	標本	属性	指向	項目	標本	属性	指向
<u>#2.3c 家庭に満足か</u>				<u>#3.2b 「宗教心」は大切か</u>			
満足	41.6	38.5	36.5	大切	68.8	65.7	62.0
やや満足	46.2	46.0	47.9	大切でない	18.6	21.0	22.5
やや不満	10.2	12.9	13.3	その他	2.0	2.2	2.5
不満	1.5	1.7	1.6	D.K.	10.6	11.1	13.0
その他	0.0	0.0	0.0	<u>#5.6b つとめたい会社</u>			
D.K.	0.5	0.8	0.8	給料が多い会社	43.5	44.9	47.3
<u>#2.10 幸福かためになることか</u>				家族的な雰囲気	53.5	52.1	48.5
しあわせにくらす	34.5	36.6	38.5	その他	0.7	0.5	0.4
世の中のためになる	62.9	60.6	57.3	D.K.	2.3	2.6	3.8
その他	0.5	0.4	0.5	<u>#7.25 お金と仕事</u>			
D.K.	2.1	2.4	3.7	仕事がなければまらない	76.4	73.6	71.4
<u>#2.11 好きなくらし方か人のためか</u>				仕事がなくともよい	21.6	23.9	25.8
自分の好きなこと	39.6	41.9	45.6	その他	0.4	0.4	0.4
人のためになること	54.1	51.8	47.6	D.K.	1.5	2.0	2.4
その他	2.6	2.4	2.2	<u>#8.6 選挙への関心</u>			
D.K.	3.7	3.9	4.6	なにをおいても投票	40.3	36.7	33.7
<u>#2.12b スキがあれば利用されるか</u>				なるべく投票	47.8	48.9	49.0
利用しようとしている	27.5	29.4	30.9	あまり投票する気がない	6.0	6.5	8.0
そんなことはない	62.4	60.2	57.1	ほとんど投票しない	5.4	7.2	8.4
その他	2.9	2.7	3.0	その他	0.0	0.0	0.0
D.K.	7.2	7.6	9.0	D.K.	0.5	0.6	0.9
<u>#2.12c 人は信頼できるか</u>				<u>#8.9 社会不満の表し方</u>			
信頼できる	30.2	28.2	23.8	選挙で考慮	55.0	50.4	47.5
用心する	63.6	65.4	69.6	合法的に	10.4	9.2	9.3
その他	1.9	1.7	1.3	場合により非法的手段	4.1	5.3	5.6
D.K.	4.3	4.7	5.3	なにもしない	28.4	32.6	34.5
<u>#2.35 ボランティア活動</u>				その他	0.2	0.2	0.2
現在している	15.9	13.6	12.8	D.K.	2.0	2.2	2.8
過去にした	29.2	26.2	25.2	<u>#9.12b 日本の「芸術」</u>			
そのうちする	26.0	24.8	23.0	非常によい	12.6	12.3	12.8
今後もしない	27.7	34.1	36.8	ややよい	59.0	55.6	53.5
その他	0.1	0.2	0.2	ややわるい	17.5	19.0	19.3
D.K.	1.1	1.1	1.9	非常にわるい	2.2	2.3	2.2
<u>#3.6 宗教か科学か</u>				その他	0.1	0.1	0.1
宗教否定	9.1	10.6	12.6	D.K.	8.6	10.7	12.1
宗教科学協力	48.5	44.3	39.3	<u>#9.16 国際貢献</u>			
宗教のみ	2.4	3.1	3.3	外国助ける	38.8	37.3	32.3
両方否定	31.5	32.3	32.7	生活水準上げる	55.9	57.6	61.4
その他	0.6	0.7	0.5	その他	1.0	0.8	0.8
D.K.	7.9	9.0	11.7	D.K.	4.2	4.3	5.4

注) 標本は標本割合, 属性は属性変数のみを用いたキャリブレーション推定値, 指向は調査への指向性変数を追加したキャリブレーション推定値.

総選挙があるとき ‘なにをおいても投票’ する (#8.6) 割合が減り, 社会に対して不満があっても ‘なにもしない’ (#8.9) 割合が増える. 補正值では “#2.3c 家庭に満足か” の不満が増え, “#9.12b 日本の「芸術」” では否定的な意見が増えるようである.

土屋 (2005, p. 41) は事後郵送調査の未返送者の特徴として『今現在の自分の身近な生活を満足させることに関心が向いており, 濃密な人間関係や, 社会一般に対してはより無関心なようである. しかしだからと言って, 未返送者が生活や社会に必ずしも満足しているわけではなさ

そう』という特性を見出している。補正値は標本割合と比べてまさにこのような方向に動いている。上記の未返送者の特徴が調査不能者の特徴として当てはまるのであれば、母集団全体の推定値として、補正値は標本割合よりも妥当であると考えてよいであろう。なお土屋(2005)では、未返送者の特徴として他に D.K. の多さも指摘している。また Couper (1997) や Triplett et al. (1996) も調査に対して消極的な人は D.K. が多くなることを認めている。表 7 に示す項目では、標本割合よりも補正値の D.K. は大きくなっており、この点でも補正値の方が標本割合より妥当であると考えられる。

4. 考察と今後の課題

調査不能バイアスの補正方法に関する研究では、調査不能ゆえに真の状態が分からず、断定的な結論は下せないのが常である。本稿においても標本割合 \hat{p}_k 、属性変数だけを用いたキャリブレーション推定値 \hat{p}_k^c 、指向性変数も加えたキャリブレーション推定値 \hat{p}_k^e のいずれが真値に最も近いのかは知りようがない。しかしこれまでの研究から窺える調査不能者の特徴に鑑みれば、標本割合と比べたキャリブレーション推定値の変化の方向は十分納得できるものである。また例えば土屋(2006)のシミュレーションによれば、キャリブレーションを行っても誤差は全体で 1~2 割程度しか縮小しない。属性変数だけでは誤差は縮小しきれておらず、指向性変数を加えたキャリブレーション推定値 \hat{p}_k^e の誤差が最も小さい可能性は高い。ただし指向性変数による推定値であっても、調査不能によるバイアスは依然として大きいおそれがある点には注意しなければならない。

指向性変数を用いると、標本割合とは 10 ポイント近く推定値が変わるカテゴリも見られた。例えば“#2.35 ボランティア活動”の‘今後もしない’などである。10 ポイントも異なれば、結果の見方に大きな影響を及ぼすことも多いであろう。調査結果としての公表値は仮に標本割合 \hat{p}_k を採用するとしても、回収率が低い調査では指向性変数を加えたキャリブレーション推定値も求め、結果の解釈に当たっては両者を総合的に勘案する必要がある。

本稿ではウェイトのキャリブレーションを二段階で行う際に、不在と拒否のそれぞれに対応した指向性変数を用いることを提案した。このうち不在に応じた指向性変数の一つ“#1.80 休日は在宅か”は、回収標本についてしか回答が得られていない。不在の調整に用いる指向性変数は本来、在宅標本全体から回答が得られている必要がある。本稿では在宅標本における分布は回収標本における分布と等しいと仮定してキャリブレーションに用いた。この仮定が全く非現実的であるとは考えにくいだが、より適切な変数を今後も探る必要がある。また山内・米倉(2002)は、本来「不在」に分類されるべきにもかかわらず、「拒否」と分類されてしまう場合があることを指摘している。二段階でキャリブレーションを行う前提として、不能理由を調査員が丁寧に聞き取ると同時に、その客観的な分類ができるよう分類基準の標準化が求められる。

キャリブレーションに用いる母集団値は、可能な限り正確な値を用いる必要がある。本稿では表 3 においてやや古い値を用いたり、対象年齢以外の人を含んだ値を用いたりした。その理由の一つは、適切な区分で国勢調査の結果が集計・公表されていないためである。しかしこの点に関しては、新統計法の下で今後オーダーメード集計が可能となれば、自然と解決されるであろう。

参 考 文 献

- Couper, M. P. (1997). Survey introductions and data quality, *Public Opinion Quarterly*, **61**, 317-338.
 Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling, *Journal of the*

- American Statistical Association*, **87**, 376–382.
- Deville, J.-C., Särndal, C.-E. and Sautory, O. (1993). Generalized raking procedures in survey sampling, *Journal of the American Statistical Association*, **88**, 1013–1020.
- Folsom, R. E. (1991). Exponential and logistic weight adjustment for sampling and nonresponse error reduction, *Proceedings of the Social Statistics Section, American Statistical Association*, 197–202.
- Folsom, R. E. and Singh, A. C. (2000). The generalized exponential model for sampling weight calibration for extreme values, nonresponse, and poststratification, *Proceedings of the Survey Research Methods Section, American Statistical Association*, 598–603.
- Groves, R. M. and Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*, John Wiley & Sons, New York.
- Holt, D. and Elliot, D. (1991). Methods of weighting for unit non-response, *The Statistician*, **40**, 333–342.
- 星野崇宏, 繁柁算男(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, **31**(1), 43–61.
- 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.
- 中村 隆, 前田忠彦, 土屋隆裕, 松本 渉(2009). 国民性の研究 第12次全国調査 — 2008年全国調査一, 統計数理研究所研究レポート, No. 99.
- Oh, H. L. and Scheuren, F. J. (1983). Weighting adjustment for unit nonresponse, *Incomplete Data in Sample Surveys*, Volume 2 (eds. W. G. Madow, I. Olkin and D. B. Rubin), 143–184, Academic Press, New York.
- 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.
- 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.
- 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, No. 92.
- Särndal, C.-E. and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*, John Wiley & Sons, Chichester.
- Stinchcombe, A. L., Jones, C. and Sheatsley, P. (1981). Nonresponse bias for attitude questions, *The Public Opinion Quarterly*, **45**, 359–375.
- Triplet, 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.
- 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, **53**(1), 35–56.
- 土屋隆裕(2006). 「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 —, 日本統計学会誌, **36**(1), 1–23.
- 土屋隆裕(2009). 『概説標本調査法』, 朝倉書店, 東京.
- 山内利香, 米倉 律(2002). 調査不能の現状と課題 ～ 「あなたから見た世論調査2」から ～, 放送研究と調査, 2002年8月号, 110–125.

Two-stage Non-response Bias Adjustment
Using Variables on the Survey-orienting Character
for the Survey on the Japanese National Character

Takahiro Tsuchiya

The Institute of Statistical Mathematics

The response rate of surveys on the Japanese national character has declined to 51.6%, which is the lowest ever. Although simple proportions among respondents have hitherto been used as population estimators, they will certainly exhibit serious biases in some question items. This paper attempts to reduce these biases by adjusting the estimation weights. The two main reasons for non-response are absence and refusal. Hence, the weights were calibrated in two stages that accounted for each reason. They were calibrated using not only demographic variables, but also variables that measure respondents' attitudes toward the survey in each stage. We consider the latter variables to be those on the survey-orienting character used in this paper. The calibrated estimates differ by ten percentage points at most from the simple proportions. The differences were larger in the calibrated estimates that used variables on the survey-orienting character than in the calibrated estimates that used only the demographic variables. Although we cannot conclude that the estimates are totally free from biases, the calibrated estimates appear to be reasonable, considering the known characteristics of the nonrespondents.