

日本における国際成人力調査のウェイト調整

Weighting Adjustment Applied to the Japanese Data of the PIAAC Survey

土屋 隆裕*

TSUCHIYA Takahiro

Abstract

Response rate of the PIAAC survey in Japan was approximately 50%. Weighting adjustment was indispensable in order to compensate for the nonresponse and frame imperfection. This paper describes the details of the weighting method applied to the Japanese data. First, the sampling design was presented in the introduction, which was considered to compute the base weight. Unfortunately due to the Great East Japan Earthquake, we couldn't conduct the survey in some areas. We addressed this problem by combining some strata. Next, five steps of weighting adjustment were shown one by one: unknown eligibility adjustment, nonliteracy-related nonresponse adjustment, literacy-related nonresponse adjustment, trimming and calibration. Finally, the results of the weighting adjustment were evaluated, and some ideas for improving the weighting adjustment toward future surveys were discussed.

1. 本論文の目的と日本における国際成人力調査の標本設計

1.1 本論文の目的

国際成人力調査は、各国に居住する16歳以上65歳以下の人を母集団とし、標本調査によって実施されている。標本調査の結果から母集団の推定を行うためには、標本抽出デザインを反映したウェイトを用いて結果の集計を行う必要がある。また、そのウェイトは、標本抽出枠の不完全さや、未回収が生じることによって起こり得る結果の偏りを補正する役割も担うものでなければならない。ウェイトの算出方法は、国際コンソーシアムが定めた技術基準においてその原則が提示されているものの、実際の適用に当たっては各国の状況に応じて最適な手法を選択していくことが求められている。

日本における標本抽出の方法やウェイト算出の方法については、その概略が既に他で紹介されている(靱井・土屋, 2012; 国立教育政策研究所, 2013)。本論文の目的は、実際に行った標本抽出の方法と、ウェイト調整の方法をさらに詳細に紹介することである。仮に次回の国際成人力調査が実施されれば、今回得られた知見を生かして、よりよい標本抽出デザインやウェイト調整法を開発することが可能となるであろうし、また他の同様の調査の設計や分析に当たっても大いに資するであろうと考えられるからである。

* 情報・システム研究機構統計数理研究所：〒190-8562 東京都立川市緑町 10-3

1.2 国際成人力調査の日本における標本設計

まずこの節では、日本における標本設計についてまとめておく。国際成人力調査では、各国の状況に応じて、標本の抽出枠として住民の名簿を用いるか、住居のリストを用いることとされている。日本では、住民名簿である住民基本台帳を抽出枠とした。抽出した全ての対象者の性と年齢が調査実施前にあらかじめ分かるからである。この点は、後にウェイト調整を行う際に非常に大きな利点となる。また日本において、住民を対象とした標本調査の多くは住民基本台帳を抽出枠としており、調査実施会社も実査に慣れていると期待できたことも理由の一つである。なお、目標母集団には含まれる外国人が住民基本台帳には記載されていないという不完全さはあるものの、その割合はごくわずかであり、住民名簿としての精度は十分に高いと考えられた。

抽出標本は、全ての対象者の包含確率が等しい自己加重標本（土屋，2007）であることが技術基準において要求された。そこで日本では層化確率比例二段抽出法を用いて標本を選ぶこととした。まず層化は表1に示すとおり、人口規模と地域で行った。次に第一次抽出単位は町丁字とし、全体の抽出数は500として、各層の抽出数は15歳以上64歳以下人口に比例割当により定めた。目標母集団の年齢は16歳以上65歳以下であるにもかかわらず、割当に用いる年齢を15歳以上64歳以下としたのは、これらの抽出作業を、調査実施のほぼ1年前に当たる2010年秋から2011年初頭にかけて行ったためである。

第二次抽出単位は住民個人とし、2011年4月以降に住民基本台帳から抽出を行う予定であった。しかし2011年3月11日に発生した東北地方太平洋沖地震に伴う東日本大震災によって、既に抽出された地点のうち一部は調査の実施が不可能となった。2011年夏からの調査実施というスケジュールは、日本だけが遅れることができないため変更することができず、また標本設計を再度行い、地点を再抽出することも時間的に困難であった。そこで国際コンソーシアムと協議し、予備調査結果などのデータも用いて検討した結果、以下のような対応策をとることとした。つまり、地点の再抽出は行わず既に抽出された地点を用いること、災害救助法適用地域に含まれる41地点は調査を実施しないこと、調査を実施しない地点が特に東北に多いことから、人口規模ごとに東北と関東の層を合併し、調査を実施する地点は合併層から抽出されたものとして扱うこと、目標母集団には被災地も含めること、一部の地域を調査しないことにより生じ得る偏りは調査実施後にウェイトによって調整すること、である。

以上の方針に基づき、各層で抽出する対象者数は、表1に示すとおり合併後の層の15歳以上64歳以下人口に比例割当することとした。さらに全体で抽出する対象者数は以下の方法で定めることとした。まず技術基準では5,000人以上の回収および50%以上の回収率が要求されており、その要求を満たせるだけの対象者を抽出しておく必要がある。しかし回収率をあらかじめ精確に予測することは難しい。5,000人を確実に確保するため、回収率を過度に低めに見積もり、あまりに多くの対象者に対して調査を実施すると、コストがかさむ上に調査員への負担も大きくなり、回収データの質が低下するおそれがある。そこで回収率は少なくとも3割弱は確保できると予想し、まず住民基本台帳からの抽出は全体の対象者数を18,000として行った。次に、抽出した対象者のうちの10,000人に対して調査を実施する。回収率は50%以上が要求されているからである。そして調査実施状況を見ながら、必要に応じて対象者を追加していくこととした。結果として追加は1,000人とどまり、全体で11,000人を調査することとなった。なお全体の対象者数が変わっても、各層への割当数は常に比例割当で求めた。以降ではこの11,000人を抽出標本として扱う。表1では、地点当たり抽出対象者数が合併層で多いが、これは合併層の地点数は比例割当よりも少ないのに対し、抽

出対象者数は比例割当となっているためである。

表 1 標本抽出デザイン

層番号	人口規模	地域	母集団 15歳以上64歳 以下人口	抽出 地点数	調査 地点数	層合併後 調査地点数	抽出 対象者数	地点当たり 抽出対象者数
1:		東京都区部	5,834,549	36	36	36	788	21.889
2:		北海道	1,287,054	8	8	8	174	21.750
3:		東北	684,374	4	0			
4:		関東	4,232,187	26	26	26	664	25.538
5:	人口100万人以上の市	中部	1,468,385	9	9	9	198	22.000
6:		近畿	3,544,464	22	22	22	479	21.773
7:		中国・四国	760,292	5	5	5	102	20.400
8:		九州	957,871	6	6	6	129	21.500
9:		東北	1,520,154	9	3			
10:		関東	7,507,295	46	42	45	1,219	27.089
11:	人口20万人以上100万人未満の市	中部	4,628,964	28	28	28	625	22.321
12:		近畿	4,972,765	31	31	31	672	21.677
13:		中国・四国	2,329,186	14	14	14	315	22.500
14:		九州	3,018,641	19	19	19	408	21.474
15:	人口10万人以上100万人未満の市	北海道	975,593	6	6	6	132	22.000
16:		東北	790,328	5	3			
17:		関東	5,111,913	31	25	28	797	28.464
18:	人口10万人以上20万人未満の市	中部	2,424,413	15	15	15	327	21.800
19:		近畿	2,188,442	13	13	13	296	22.769
20:		中国・四国	1,623,172	10	10	10	219	21.900
21:		九州	1,304,369	8	8	8	176	22.000
22:		北海道	623,564	4	4	4	84	21.000
23:		東北	1,705,589	10	4			
24:		関東	3,741,706	23	16	20	736	36.800
25:	人口10万人未満の市	中部	3,875,500	24	24	24	523	21.792
26:		近畿	2,811,425	17	17	17	380	22.353
27:		中国・四国	1,634,550	10	10	10	221	22.100
28:		九州	2,584,084	16	16	16	349	21.813
29:		北海道	631,738	4	4	4	85	21.250
30:		東北	1,129,063	7	2			
31:		関東	1,348,231	8	7	9	335	37.222
32:	町村	中部	1,270,176	8	8	8	171	21.375
33:		近畿	888,727	5	5	5	120	24.000
34:		中国・四国	760,457	5	5	5	102	20.400
35:		九州	1,285,796	8	8	8	174	21.750
	合計		81,455,017	500	459	459	11,000	

1.3 回収状況

表 2 は 11,000 人の調査結果の状況を分類したものである。どの分類を「回収」として扱うかは、回収率を計算するときと、ウェイトを算出したり結果データの分析を行ったりするときとは異なる。回収率を求めるときには処理コード 07 (言語上の問題) や処理コード 90 (技術上の問題) も「回収」に含めるが、ウェイトを算出するときには処理コード 01 (完了) のみが「回収」となる。本稿ではウェイト調整の問題を扱うため、表中の「ウェイト算出・データ分析のための分類」を用いることとし、処理コード 01 (完了) のみを「回収」として扱う。

調査の結果回収とされ、データ分析に用いられるのは 5,173 人であり、抽出された 11,000 人の半分に満たない。逆に言えば抽出された対象者の半分以上が集計に用いられず、回収された 5,173 人

を単純に集計しただけでは偏りが生じるおそれがある。例えば抽出された 11,000 人のうち男性は 50.4% (5,542 人) であるが、回収となった 5,173 人の中では 47.7% (2,468 人) である。そこで、本稿で述べるウェイトを用いることによって男性の重みを拡大し、結果の偏りを軽減する必要がある。

表 2 調査結果の状況の分類

処理コード		回収率算出のための分類	ウェイト算出・データ分析のための分類	人数	
01	完了	回収	R(回収)	5,173	
03	中断	未回収	NR(リテラシー関連ではない未回収)	0	
04	拒否(対象者による)	未回収	NR(リテラシー関連ではない未回収)	2,937	
05	拒否(対象者以外の者による)	未回収	NR(リテラシー関連ではない未回収)	1,153	
07	言語上の問題	回収	性・年齢既知	L1(リテラシー関連未回収)	6
			性・年齢未知	L2(リテラシー関連未回収)	0
08	非識字	回収	性・年齢既知	L1(リテラシー関連未回収)	2
			性・年齢未知	L2(リテラシー関連未回収)	0
09	学習・精神障害	回収	性・年齢既知	L1(リテラシー関連未回収)	97
			性・年齢未知	L2(リテラシー関連未回収)	0
12	聴覚障害	除外	D(障害)	7	
13	視覚障害	除外	D(障害)	13	
14	発語障害	未回収	NR(リテラシー関連ではない未回収)	3	
15	身体障害	除外	D(障害)	36	
16	その他の障害	除外	D(障害)	34	
17	その他	未回収	NR(リテラシー関連ではない未回収)	160	
18	死亡	除外	I(不適格)	10	
21	非接触	未回収	NR(リテラシー関連ではない未回収)	460	
24	一時不在	未回収	NR(リテラシー関連ではない未回収)	428	
	移転	未回収・除外			U(適格性未知)
25	不適格	除外	I(不適格)	85	
90	技術上の問題	回収	NR(リテラシー関連ではない未回収)	3	
合計				11,000	

2. ウェイト調整の方法

2.1 ウェイト調整の概要

ウェイトを用いるということは、例えば変数 Y の第 l 対象者の値を Y_l としたとき、変数 Y の平均の推定に当たって、単なる標本平均 $\bar{Y} = n^{-1} \sum_l Y_l$ を用いるのではなく、第 l 対象者にはウェイト FW_l を与え、このウェイトによる加重平均

$$(2.1) \quad \hat{\mu}_Y = \frac{\sum_l FW_l Y_l}{\sum_l FW_l}$$

を用いるということである。各対象者のウェイト FW_l は、技術基準において、(2.2)式に示すように標本抽出デザインを反映した基礎ウェイト W_l に対し係数 F_{1l} から F_{6l} を順に掛けていくことで求めることとされている。

$$(2.2) \quad FW_l = W_l F_{1l} F_{3l} F_{4l} F_{5l} F_{6l}$$

各係数の意味および計算方法は以下のとおりである。

$$\begin{aligned}
(2.3) \quad & F_{1l}: \text{適格性未知の調整用係数} \\
& = \begin{cases} \frac{S_R + S_{NR} + S_{L1} + S_{L2} + S_D + S_I + S_U}{S_R + S_{NR} + S_{L1} + S_{L2} + S_D + S_I} & l \in I \\ \frac{S_R + S_{NR} + S_{L1} + S_{L2} + S_D}{S_R + S_{NR} + S_{L1} + S_{L2} + S_D + S_I} & l \in U \\ 1 & \square \quad \square \quad l \in R, NR, L1, L2, D \end{cases} \\
(2.4) \quad & F_{3l}: \text{リテラシー関連ではない未回収の調整係数} \\
& = \begin{cases} 1 & \square \quad l \in L1, L2, I \\ \frac{S_R + S_{NR} + S_D + S_U}{S_R} & l \in R \\ 0 & \square \quad l \in NR, D, U \end{cases} \\
(2.5) \quad & F_{4l}: \text{リテラシー関連の未回収の調整係数} \\
& = \begin{cases} 1 & \square \quad l \in R, I \\ \frac{S_{L1} + S_{L2}}{S_{L1}} & l \in L1 \\ 0 & \square \quad l \in L2 \end{cases} \\
(2.6) \quad & F_{5l}: \text{トリミング用係数} \\
& = \begin{cases} 1 & \square \quad W_l F_{1l} F_{3l} F_{4l} \leq \text{cutoff} \\ \frac{\text{cutoff}}{W_l F_{1l} F_{3l} F_{4l}} & W_l F_{1l} F_{3l} F_{4l} > \text{cutoff} \end{cases} \\
(2.7) \quad & F_{6l}: \text{キャリブレーション用係数}
\end{aligned}$$

各式における S_* は、表 2 に示される分類*における、それよりも一つ前の段階までのウェイト合計を表す。例えば F_{1l} における S_R とは、表 2 においてウェイト算出のための分類が **R** (回収) とされた対象者、すなわち処理コード **01** とされた対象者のウェイト W_l の合計を意味する。また F_{3l} における S_R とは、**R** (回収) の対象者のウェイト $W_l F_{1l}$ の合計を意味する。したがって例えば係数 F_{3l} を乗じるということは、**NR** (リテラシー関連ではない未回収) と **D** (障害)、**U** (適格性未知) の対象者のウェイトは $W_l F_{1l} F_{3l} = 0$ として集計から除外し、**R** (回収) の対象者のウェイトをその分拡大するというを意味する。

ただし上記の係数のうち F_{1l} から F_{4l} については、各係数を求めるためにそれぞれ対象者をその属性によってウェイト調整群に分割し、ウェイト調整群ごとに係数を求める。これは表 2 の各処理コードに分類される対象者の割合が属性によって異なるからである。先に挙げた例で言えば、女性よりも男性の方が **R** (回収) となる割合は低い。そこで男性のウェイトは女性のウェイトよりもさらに拡大することで、集計結果の偏りを軽減しようというのである。したがって上記の S_* の説明をより正確に述べれば、例えば F_{1l} における S_R とは、表 2 においてウェイト算出のための分類が **R** (回収) とされた対象者のうち、あるウェイト調整群に属する対象者のウェイト W_l の合計を意味することとなる。

2.2 ウェイト調整群のための変数選択

係数を求めるためのウェイト調整群への分割は、対象者の属性をもとに行われる。まず F_{1i} を求めるためのウェイト調整群は、最も基本的な属性である性と 10 歳刻み年齢層とを組み合わせたものを用いることとした。これは表 2 から分かるとおり、I (不適格) とされた対象者がわずか 85 名に過ぎず、多数のウェイト調整群に分割することは適切ではないと判断したためである。また F_{4i} については、L2 (リテラシー関連未回収、性・年齢未知) とされた対象者がいなかったため、結果的に全ての対象者が $F_{4i} = 1$ となっており、ウェイト調整群への分割は不要であった。

残る F_{3i} を求めるためのウェイト調整群への分割は、調査で目的とする変数すなわち成人力との相関が高く、かつ回収・未回収との関連も強い変数を用いることが望ましい (Lumley, 2010, pp.192)。成人力との相関は高いが、回収・未回収とは関連がない変数ではウェイトを調整しても集計結果は変わらず、回収・未回収とは関連していても成人力との相関がない変数では、推定値の分散を拡大させるだけだからである。

そこでまずウェイト調整群を構成するための属性変数の候補を多数用意し、その中から適切な変数を選び出すこととした。候補となる変数は、抽出された全ての対象者について値が得られなければならない。対象者の個人レベルで値が分かる変数は、住民基本台帳から得られる性と年齢、および調査員の観察に基づく対象者の住居形態のみである。そこで、平成 22 年度国勢調査の結果を用いて地域レベルでの変数を対象者に付与することとした。標本の第一次抽出単位は町丁字であるため、本来は町丁字レベルでの国勢調査結果を利用するのがよいと考えられる。しかし、ウェイトの作成作業は 2012 年春から夏にかけて行う必要があったのに対し、国勢調査の小地域集計結果が全て公表されたのは 2012 年 12 月 11 日であったため、町丁字レベルでの結果は利用できず、市区町村レベルでの集計結果を用いることとした。表 3 に示すのは、候補として用いた変数の一覧である。

次に表 3 に示す変数の中から成人力との相関が高い変数を選び出した。ただしウェイトの作成を行う時点で、各対象者の成人力関連の変数値は未だ国際コンソーシアムから提示されておらず、相関係数を直接求めることは不可能であった。そこで国際コンソーシアムから示された方法に従い、学歴や年齢等から成人力関連の代理変数を作成し、この代理変数を各候補変数に回帰させたときの決定係数を求めた。表 3 の右列に示すのは、その決定係数の値である。代理変数との散布図も確認しながら、最終的に変数名に星印のついた 17 変数を選び出した。選び出された変数のうち連続量の変数は、分布状況を見ながら 5 段階にカテゴリ化した。各カテゴリの詳細は紙幅の都合から割愛する。

2.3 未回収の調整のためのウェイト調整群への分割

次に、R (回収) かあるいは NR+U+D (リテラシー関連ではない未回収、適格性未知、障碍) を基準変数とし、17 変数を用いて分類木を作成した。得られた分類木は図 1 のとおりである。回収・未回収と最も関連しているのは SERVICE (第 3 次産業の割合) であり、第 3 次産業就業者の割合が 75% 未満の地域 (図中では SERVICE=abc と表記) では回収率は 52.76%、75% 以上の地域 (図中では SERVICE=de と表記) では 42.9% であった。さらに第 3 次産業就業者の割合が高い地域のうち、一戸建に居住する対象者 (図中では BUILD=a と表記) の回収率は 55.28%、共同住宅などそれ以外の対象者 (図中では BUILD=bcd と表記) の回収率は 44.52% であった。このようにして、L1、L2、I の対象者を除く 10,800 人の対象者を全部で 17 のウェイト調整群に分割した。なお図 1 において 5 番目と 10 番目のウェイト調整群について複数のグループをまとめているのは、ウェイト調整群の最

小サイズは30人とするよう国際コンソーシアムから指示があったためである。

表3 ウェイト調整群作成のための候補となる変数

種類	変数名	内容	決定係数
対象者個人属性	GENDER ★	性別	
	AGE ★	年齢	
	BUILD ★	住居形態	
教育	UNIV ★	大学卒業者割合	0.0226
	EDU2	高校卒業者に対する大学卒業者の比	0.0214
	COLG ★	中学・高校卒業者に対する短大・大学卒業者の比	0.0219
従業上の地位	EMP1	総数に対する派遣社員の割合	0.0098
	EMP2	総数に対する家族従業者の割合	0.0111
	EMP3	雇用者に対する派遣社員の割合	0.0115
	TMPE ★	正規の職員に対する派遣社員の比	0.0099
	EMP5	正規の職員に対する家族従業者の比	0.0102
	EMP6	自営業と家族従業者に対する派遣社員の比	0.0104
	EMP7	自営業と家族従業者に対する役員と派遣社員の比	0.0147
	EXEC ★	自営業と家族従業者と内職に対する役員と派遣社員の比	0.0148
	EMP9	総数に対する雇用者の割合	0.0002
労働力状態	LABOR ★	人口に対する労働力人口の割合	0.0126
	LBF2	人口に対する非労働力人口の割合	0.0001
	LBF3	労働力人口に対する非労働力人口の割合	0.0028
	LBF4	人口に対する就業者の割合	0.0100
産業	IND1	第1次産業の割合	0.0114
	IND2	第2次産業の割合	0.0033
	SERVICE ★	第3次産業の割合	0.0104
	IND4	A農業、林業の割合	0.0102
	IND5	B漁業の割合	0.0031
	IND6	C鉱業、採石業、砂利採取業の割合	0.0024
	IND7	D建設業の割合	0.0103
	IND8	E製造業の割合	0.0010
	IND9	F電気・ガス・熱供給・水道業の割合	0.0004
	IND10	G情報通信業の割合	0.0177
	IND11	H運輸業、郵便業の割合	0.0001
	IND12	I卸売業、小売業の割合	0.0011
	IND13	J金融業、保険業の割合	0.0131
	IND14	K不動産業、物品賃貸業の割合	0.0182
	IND15	L学術研究、専門・技術サービス業の割合	0.0203
	IND16	M宿泊業、飲食サービス業の割合	0.0005
	IND17	N生活関連サービス業の割合	0.0004
	IND18	O教育、学習支援業の割合	0.0026
	IND19	P医療、福祉の割合	0.0053
	IND20	Q複合サービス事業の割合	0.0136
	IND21	Rサービス業の割合	0.0044
	IND22	S公務の割合	0.0049
	INFO ★	情報通信業、金融業、保険業、不動産業、物品賃貸業、学術研究、専門・技術サービス業の割合	0.0201
地域	REGION7	北海道、東北、関東、中部、近畿、中国・四国、九州	
	TL2 ★	JPA、JPB、JPC、JPD、JPE、JPF、JPG、JPH、JPI、JPJ注)	
	CITY ★	人口規模	
	POP.D ★	人口密度	0.0118
世帯	HHS1	6人以下世帯に対する1人世帯の割合	0.0066
	H.ONE ★	全世帯に対する1人世帯の割合	0.0069
	HHS3	人口に対する世帯数の比	0.0087
	H.SIZE ★	世帯数に対する人口の比	0.0082
住居の面積	H.AREA ★	全住居に対する100平米以上の住居の割合	0.0087
	HS2	全住居に対する150平米以上の住居の割合	0.0064
年齢構成	AGED1	生産年齢人口の割合	0.0115
	AGED2	65歳以上人口の割合	0.0085
	AGED3	15歳未満あるいは65歳以上人口の割合	0.0111
	AGED4	65歳未満人口に対する65歳以上人口の比	0.0085

注)OECDによる地域区分であり、各記号については <http://www.oecd.org/regional/regional-policy/43428422.pdf> あるいは表4を参照のこと

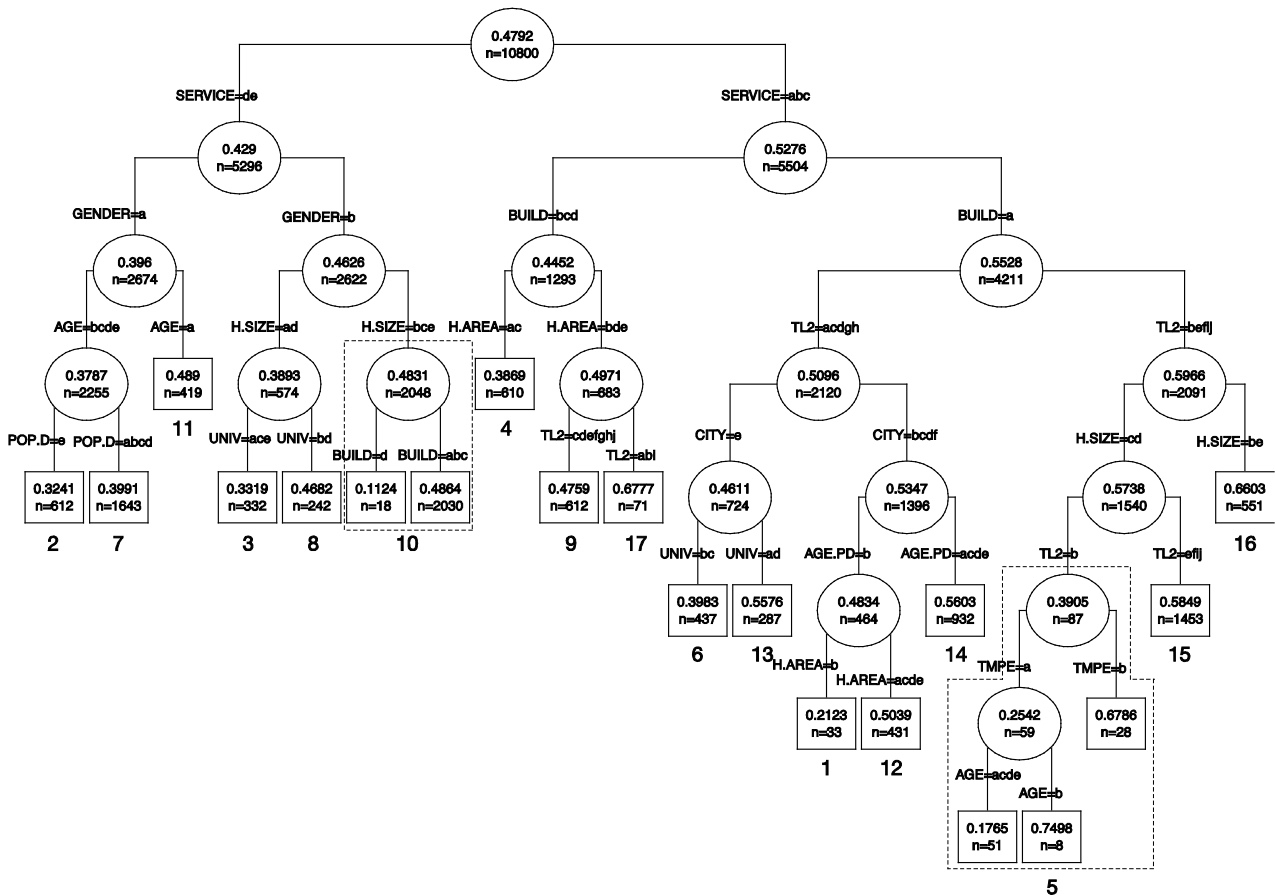


図 1. 回収・未回収を基準変数とした分類木とウェイト調整群への分割

2.4 トリミングおよび母集団情報を用いたキャリブレーション

係数 F_{5l} で用いられるトリミングの基準値 (cutoff) については、国際コンソーシアムからの指示で以下のように決めることとした。すなわちまずトリミングを行わずにキャリブレーションを行って暫定的に係数 F_{6l} を求め、次に暫定的なキャリブレーション後のウェイト $W_l F_{3l} F_{4l} F_{6l}$ の変動係数 CV と中央値 M を用いて、 $\text{cutoff} = 3.5 \times \sqrt{1 + CV^2} \times M$ と定めた。この基準に従うと、日本ではトリミングは行われなかった。

係数 F_{6l} を求めるためのキャリブレーションとは、基本的な属性に関して、回収標本による推定値と母集団の真の値が一致するようウェイトを調整する方法である (Deville & Särndal, 1992; Deville, Särndal & Sautory, 1993)。仮に回収率が 100%であったとしても、標本調査であれば、例えば男女比が回収標本と母集団とで一致するとは限らない。そこでウェイトの調整によって両者を一致させ、標本誤差を低減させるのがキャリブレーションの目的の一つである。また先述のとおり、東日本大震災後に東北と関東を一つの層に合併したため、東北の標本の割合は母集団に比べ小さく、逆に関東の標本の割合は大きくなっている。この歪みは、回収率によるウェイト調整 F_{3l} では補正することはできない。そこで母集団情報と照らし合わせることで東北の回収標本には大きなウェイトを与えると同時に、関東の回収標本には小さなウェイトを与えることで、抽出標本に起因する歪みを補正することがキャリブレーションのもう一つの目的である。

キャリブレーションに用いる変数は年齢層、性、学歴、就業状態、地域の5つとした。これらは基礎的な重要な属性変数であり、成人力との相関もある上に、母集団における総数も得られるからである。なお母集団における総人口は、2011年10月1日現在の人口推計（総務省統計局）に基づく総数81,486,438人から、2008年10月1日現在の患者調査（厚生労働省）における入院患者数427,200人を引いた81,059,238人を基準人口とした。2011年10月1日現在の患者調査結果は、ウェイト作成時には未公表だったためである。就業状態に関しては平成23年10月から12月の労働力調査結果（総務省統計局）、その他の変数については平成22年国勢調査結果（総務省統計局）を利用して作成した。各変数の母集団における分布は表4の左の FW_i （母集団）の列に示すとおりである。

表4 基本属性の分布および基本属性別リテラシー平均

		割合		リテラシー平均			
		W_i	FW_i	W_i		FW_i	
		(回収標本)	(母集団)	推定値	標準誤差	推定値	標準誤差
全体		100.0%	100.0%	297.8	0.71	297.0	0.52
性別	男性	47.7%	50.3%	299.6	0.91	298.8	0.75
	女性	52.3%	49.7%	296.1	0.88	295.1	0.80
年齢層	16-25	16.1%	15.5%	300.9	1.16	299.6	1.18
	26-35	18.6%	19.4%	311.4	1.24	310.9	1.28
	36-45	23.0%	22.9%	308.2	1.01	307.7	0.95
	46-55	19.6%	19.2%	297.1	1.41	296.0	1.26
	56-65	22.8%	22.9%	274.6	1.37	273.3	1.41
学歴	中学校卒	7.0%	7.8%	252.5	2.08	254.1	2.12
	高等学校卒	36.7%	42.5%	286.7	0.84	287.5	0.83
	短大卒	20.6%	17.8%	302.4	1.03	302.8	1.11
	大学卒	25.0%	23.8%	321.2	0.84	322.0	0.85
	在学中	10.7%	8.1%	301.9	1.44	301.6	1.41
就業状態	就業者	74.9%	72.2%	298.9	0.74	298.2	0.54
	完全失業者	2.3%	3.4%	307.0	4.10	307.4	4.02
	非労働力	22.8%	24.4%	293.2	1.25	291.7	1.23
地域	JPA(北海道)	4.4%	4.2%	296.9	2.72	296.5	2.78
	JPB(青森県・岩手県・宮城県・秋田県・山形県・福島県)	4.3%	7.0%	286.6	4.69	287.3	4.40
	JPC(埼玉県・千葉県・東京都・神奈川県)	30.2%	29.2%	303.7	1.48	303.2	1.40
	JPD(茨城県・栃木県・群馬県・山梨県・長野県)	6.8%	7.7%	293.9	2.07	292.9	2.19
	JPE(新潟県・富山県・石川県・福井県)	4.6%	4.1%	297.4	2.03	297.0	2.51
	JPF(岐阜県・静岡県・愛知県・三重県)	13.3%	11.8%	297.7	1.86	298.0	1.74
	JPG(滋賀県・京都府・大阪府・兵庫県・奈良県・和歌山県)	14.9%	16.4%	299.9	1.90	298.8	1.87
	JPH(鳥取県・島根県・岡山県・広島県・山口県)	5.7%	5.6%	297.0	3.67	296.0	3.68
	JPI(徳島県・香川県・愛媛県・高知県)	3.7%	2.9%	297.4	3.75	294.7	3.64
	JPJ(福岡県・佐賀県・長崎県・熊本県・大分県・宮崎県・鹿児島県・沖縄県)	12.0%	11.1%	287.3	1.69	286.9	1.57

ウェイトのキャリブレーションには、主に3つの方法が知られている。すなわち一般化回帰推定量、レイキング比推定量、ロジット関数を用いたキャリブレーションの3つである（土屋, 2007）。一般化回帰推定量（Cassel, Särndal, & Wretman, 1976）は、キャリブレーションに用いる5つの変数の第*l*対象者の値のベクトルを x_l とすると、 F_{6l} を次式とするものである。

$$(2.8) \quad F_{6l} = 1 + (\boldsymbol{\tau}_x - \hat{\boldsymbol{\tau}}_x) \left(\sum_l W_l F_{1l} F_{3l} F_{4l} F_{5l} \mathbf{x}_l \mathbf{x}_l' \right)^{-1} \mathbf{x}_l$$

$\boldsymbol{\tau}_x$ は \mathbf{x} の母集団総計のベクトルであり、 $\hat{\boldsymbol{\tau}}_x$ は $W_l F_{1l} F_{3l} F_{4l} F_{5l}$ を用いた \mathbf{x} の母集団総計の推定値ベクトルである。レイキング比推定量 (Deming & Stephan, 1940) では F_{6l} は次式となる。

$$(2.9) \quad F_{6l} = \exp(\mathbf{x}_l' \boldsymbol{\beta})$$

$\boldsymbol{\beta}$ は、キャリブレーションに用いる 5 つの変数に関して、キャリブレーション後のウェイトを用いた推定値と母集団における値とが一致するよう求める係数ベクトルである。さらに、ロジット関数を用いたキャリブレーションでは F_{6l} は次式となる (Folsom, 1991; Folsom & Singh, 2000)。

$$(2.10) \quad F_{6l} = \frac{L(U - 1) + U(1 - L) \exp(\mathbf{A}\mathbf{x}_l' \boldsymbol{\beta})}{(U - 1) + (1 - L) \exp(\mathbf{A}\mathbf{x}_l' \boldsymbol{\beta})}$$

ただし $\mathbf{A} = (U - L) / \{(U - 1)(1 - L)\}$ であり、 U と L はそれぞれ F_{6l} の上限と下限としてあらかじめ指定しておく値である。

ウェイトのキャリブレーションを行う際には、これら 3 つの方法のいずれを採用するか、さらにロジット関数を用いるのであれば上限 U と下限 L の値をいくつにすべきかを決めなければならない。一つの考え方は、キャリブレーション後のウェイトのバラツキが小さく、極端なウェイトが与えられていない方法を選ぶというものである (土屋, 2012)。特定の対象者にのみ極端に大きなウェイトが与えられると、推定値に与える当該対象者の影響が大きくなり、標本誤差はかえって拡大してしまうからである。このようなウェイトのバラツキの大きさを表す指標として、しばしば不等加重効果 (Kish, 1965) が用いられる。

$$(2.11) \quad \text{不等加重効果 } UWE = n \frac{\sum_l FW_l^2}{(\sum_l FW_l)^2}$$

つまり各方法でそれぞれウェイト FW_l を求め、それらの中で不等加重効果が最も小さな方法を選ぶのである。

図 2 には 3 つのキャリブレーション方法をそれぞれ用いたときの最終ウェイト FW_l の分布と不等加重効果の値を示した。図中の縦の点線は抽出ウェイト W_l の平均である。一般化回帰推定量 (図中では GREG と表記) では極端に大きなウェイト FW_l は少ないが、他の方法と比べ、抽出ウェイト W_l よりも値が著しく小さなウェイト FW_l が生じてしまっている。抽出ウェイト W_l は、各対象者が標本として本来持っていた重みであり、最終的なウェイト FW_l はできる限りこれを下回らないのが望ましい。レイキング比推定量 (図中では RAKE と表記) では他の方法と比べ極端に大きなウェイトが得られてしまっている。ロジット関数では下限を $L = 0$ とし、上限を $U = 5$ から $U = 2$ まで 1 ずつ小さくしてみると、 $U = 3$ のときに不等加重効果が最も小さい。一般回帰推定量やレイキング比推定量と比べても、極端に大きなウェイトは得られておらず図 2 の中では最も望ましいウェイトの分布となっている。そこで、下限を $L = 0$ 、上限を $U = 3$ としたロジット関数を採用することとした。

なお、ロジット関数において U の値を整数とするのではなく、小数点以下まで定めることで、さらに不等加重効果を小さくすることは可能と考えられる。しかし U の値をさらに細かく定めてもウ

ウェイトの値自体は大きくは変わらないこと、国際成人力調査では標準誤差をジャックナイフ法で求めるが、 U をあまり小さくし過ぎるとジャックナイフウェイトを求められない場合が生じるおそれがあることを考慮し、上記のとおり $U = 3$ と定めた。

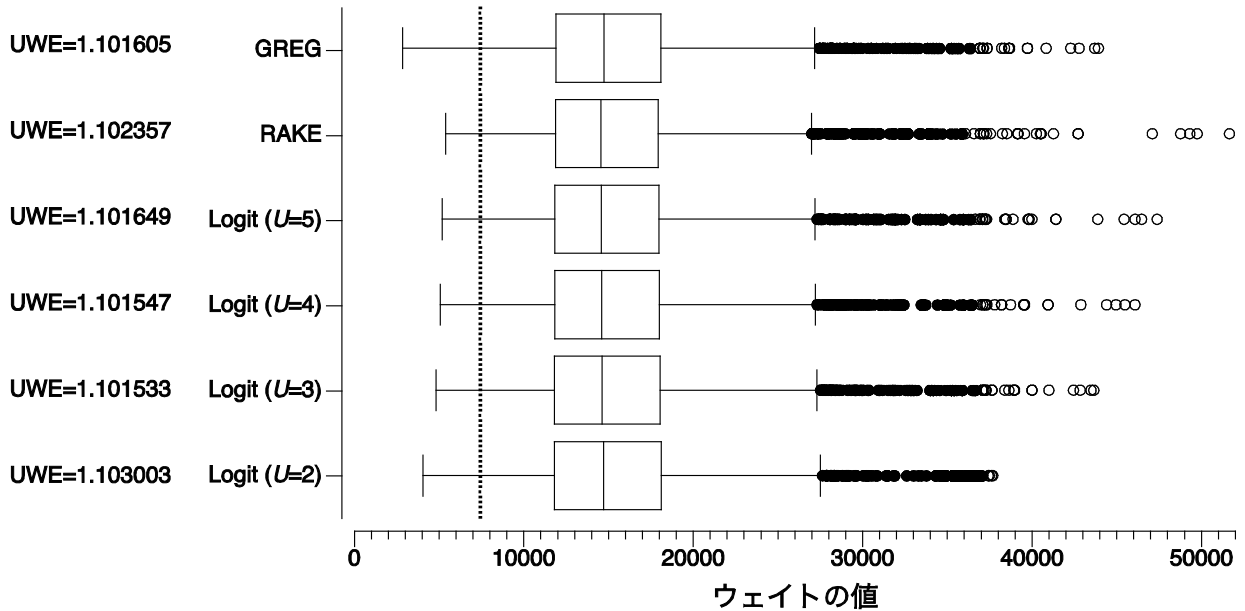


図 2. キャリブレーション結果の比較

3. ウェイト調整の結果

以上の手続きを踏んで得られた各段階のウェイト分布は図 3 に示すとおりである。自己加重標本となるよう標本の抽出を行ったため、基礎ウェイト W_l は対象者の間でほとんど一定であるが、最終的なウェイト $FW_l = W_l F_{1l} F_{3l} F_{4l} F_{5l} F_{6l}$ は、10000 を下回るものから 40000 を上回るものまで対象者の間で 4 倍以上の開きがある。

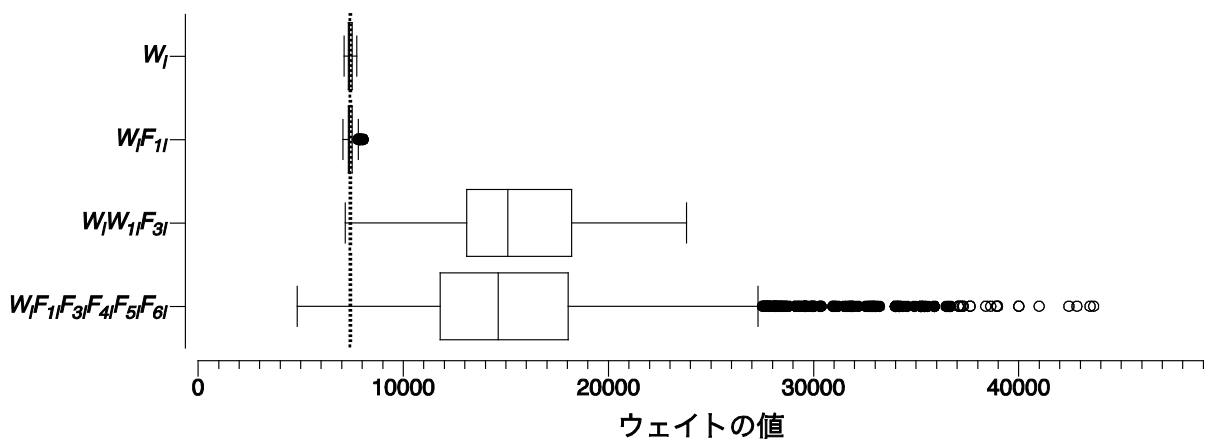


図 3. 各段階のウェイト分布

ウェイトの調整によって、成人力の平均に関する推定値がどのように変わったのかを見たのが表

4の右半分である。表4には、リテラシーに関する推算値の一つ（PVLIT1）を用いて、ウェイトの調整を行う前の基礎ウェイト W_i による平均の推定値と、最終的なウェイト FW_i による平均の推定値を属性ごとに示している。全体で見るとウェイト調整前の推定値 297.8 に対して、ウェイト調整後の推定値は 297.0 であり、やや低くなっているが、標準誤差の大きさを考慮すると著しく大きな差とは言えない。

このような結果が得られた理由としては、以下のようなことが考えられる。まず第一に、一般に母集団平均や母集団割合に関する推定では、特に標本サイズが大きいとき、ウェイトを変えても推定値自体はほとんど変わらないことが多い。平均を求めるときの分母と分子にウェイトが用いられるからである。したがって表4において、基礎ウェイト W_i と最終ウェイト FW_i の間で推定値があまり異なることはある程度予想できた結果である。

第二に、図1に示されるように、第3次産業就業者の割合が高い地域では回収率が低い。そのような地域では成人力の平均は高いため、 F_{3i} によって回収率を調整すると、全体の成人力平均の推定値は高くなる。しかし一方で、表4の左半分から分かるように、回収標本では高学歴層の割合が高い。キャリブレーション F_{6i} によって高学歴層の割合を引き下げることによって、成人力の推定値も引き下げられ、最終的には成人力の推定値はウェイト調整前と変わらなくなったというのも理由の一つと考えられる。

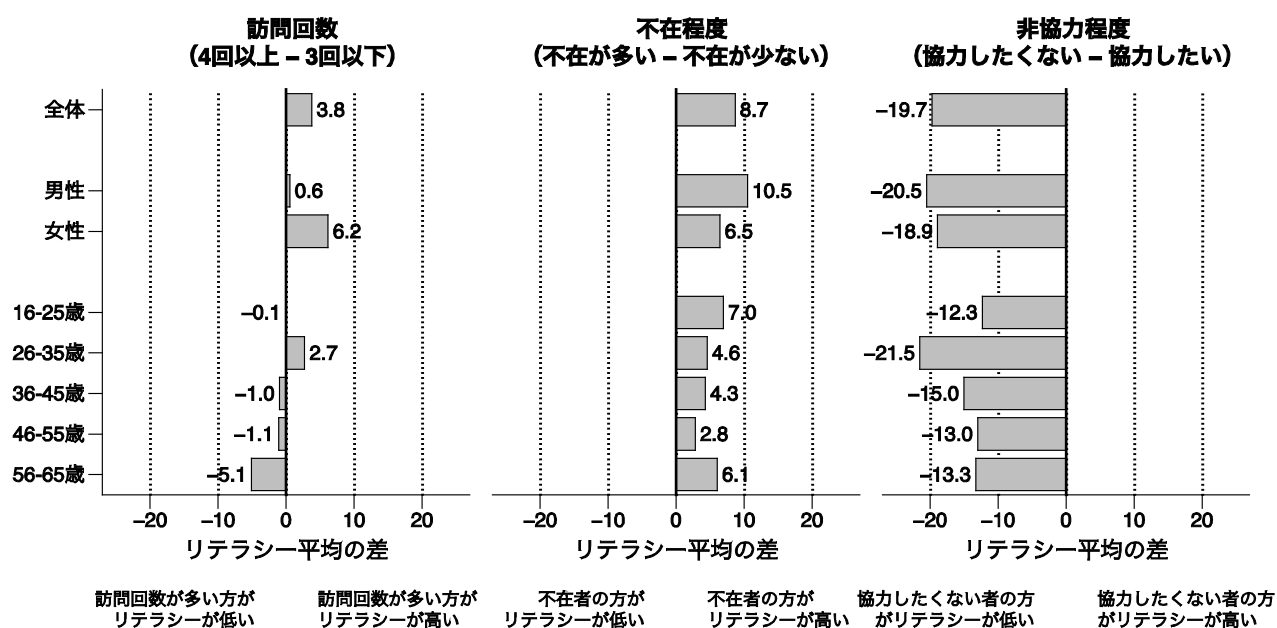


図4. 回収非困難者と比べた回収困難者のリテラシー平均

さらに別の理由としては以下のようなことも挙げられる。調査を完了した対象者の中にも、不在がちで接触までに時間を要した対象者や、最初は調査を拒否し、協力を得にくかった対象者など、回収困難な者とそうではない者がいる。図4は、そのような回収困難者のリテラシー（PVLIT1を使用）の平均を、回収非困難者の平均と比較したものである。図4の左は、訪問回数が4回以上であった完了者のリテラシー平均を、3回以下であった完了者のリテラシー平均と比べた結果であり、全体では4回以上訪問した完了者の方が平均値が3.8大きいことを表す。図4の中央は、対象者に対して『あなたは普段、ご在宅のことが多いですか、それとも外出していることが多いですか。』

と尋ね、「ほとんど在宅している (638 人)」「どちらかといえば在宅している (1,070 人)」という対象者を不在が少ない者、「どちらかといえば外出している (1,408 人)」「ほとんど外出している (2,029 人)」という対象者を不在が多い者として、リテラシー平均を比較したものである。どの性・年齢層で見ても、不在がちの完了者の方がリテラシー平均は高い。図 4 の右は、対象者に対して『仮にこのような調査がもう一度あるとしたら、あなたは協力したいと思いますか。』と尋ね、「必ず協力したい (421 人)」「協力するようにしたい (3,612 人)」と回答した対象者を協力したい者、「あまり協力したくない (225 人)」「協力したくない (892 人)」と回答した対象者を協力したくない者としてリテラシー平均を比較したものである。どの性・年齢層であっても協力したくないという対象者のリテラシー平均は低いことが分かる。

未回収者のリテラシーは結局のところ不明ではあるが、仮に回収困難者と似ているとするならば、未回収者のうち非接触者や一時不在者のリテラシー平均は回収者よりも高く、拒否者のリテラシー平均は回収者よりも低いと推測できる。したがって、仮にそれらの未回収者のデータが得られたとしても、全体の平均値は大きく変わることはないと予想される。ただし同時に、表 2 に示されるように未回収者に占める拒否者の割合は高く、かつ図 4 に示されるように協力したくない者のリテラシー平均は明らかに低いことから、ウェイト調整を行わないと全体の成人力は高めに推定されることも考えられる。したがって表 4 においてウェイト調整後のリテラシー平均が低く推定されているのは、その程度についての評価はし難いものの、その方向に関しては妥当なものと評価できよう。

4. おわりに

本稿では国際成人力調査の日本における標本設計およびウェイト調整の詳細について紹介した。未回収者の特性については、補助的なデータから様々な想像を働かせることは可能ではあるが（例えば土屋, 2005）、結局はデータが得られていない以上、「正解」を知る術はない。したがって本稿で紹介したウェイト調整の結果、目標母集団に関して「正しい」推定値が得られているのか否かを何らかの基準に照らし合わせて検証することは不可能である。しかし、ウェイトの調整は合理的な手法で行っており、得られた推計結果について妥当ではないと拒絶する理由も特に見当たらないことから、本稿でその作成手法を紹介したウェイトは適正に作成されたものであるとして受け入れてよいであろう。

同時に、次回の調査へ向けて検討・試行すべき点もいくつかある。第一に未回収の調整に用いる地域情報の精度である。地域情報としては国勢調査結果を用いることになるが、国勢調査は 5 年に一度の実施であり、調査事項が多い大規模調査は 10 年に一度しか実施されない。また、詳細な地域別の結果が公表されるまでには、調査実施後ある程度の期間を要する。そのためウェイト調整を行う時期によっては、最新ではあるが広域の情報と、古いがより小地域の情報のいずれを用いるべきかという問題が生ずる。今回は、前者の最新ではあるが広域の情報を用いたが、次回の調査に向け、後者を採用した場合の得失を検討しておく必要があるだろう。

第二に、未回収の調整に用いる変数として今回は外形的な属性変数のみを用いたが、調査に対する対象者の意識・態度変数を取り入れることで、さらに望ましいウェイト調整ができる可能性がある（土屋, 2006; 2010）。抽出された対象者のうち、容易に接触可能で、かつ調査に対して積極的に協力しようという対象者の回収率は 100%であろうと仮定し、そのような対象者の未回収調整ウェイトは $F_{3j} = 1$ とするのである（土屋, 2013; 朴・土屋, 2013）。問題は、対象者が調査に対して積極的

か否かをどのようにして判別するかである。例えば今回は、調査が終了してから『仮にこのような調査がもう一度あるとしたら、あなたは協力したいと思いますか。』と対象者に尋ね、「協力したくない」と回答した対象者のリテラシーは低いという結果が得られた。調査協りに消極的な対象者はリテラシーが低いという可能性もあるが、最初は調査依頼に積極的に応じたが、回答してみたところ難しい調査だったと感じた対象者が、二度と「協力したくない」と答えた可能性もある。調査実施前に、調査員が対象者の積極性を判断するという方法も考えられるが、判断基準を調査員間でいかに揃えるか、はたして調査員が判断できるのか課題も多い。

高い回収率が見込めない近年の調査環境においては、ウェイトの調整は調査結果の質を担保する上で不可欠である。本稿が、今後さらに優れたウェイト調整手法の開発に資することになれば幸いである。

参考文献

- Cassel, C.M., Särndal, C.-E. and Wretman, J.H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika*, 63, 615-620.
- Deming, W.E. and Stephan, F.F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11, 427-444.
- 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.
- Kish, L. (1965). *Survey Sampling*. John Wiley & Sons, New York.
- 国立教育政策研究所編 (2013) 成人スキルの国際比較 -OECD 国際成人力調査(PIAAC) 報告書一, 明石書店.
- Lumley, T. (2010). *Complex Surveys*. John Wiley & Sons, New Jersey.
- 榎井圭子・土屋隆裕 (2012) 調査技術, 国立教育政策研究所内国際成人力研究会 (編)『成人力とは何か OECD「国際成人力調査」の背景』, 明石書店, pp.133-166.
- 朴堯星・土屋隆裕 (2013) 多摩地域住民意識調査 -昭島市・小金井市郵送調査(2013)-, 統計数理研究所調査研究レポート 112.
- 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.
- 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」データへの適用一, 日本統計学会誌, 第 36 巻第 1 号, 1-23.
- 土屋隆裕 (2007) 概説標本調査法, 朝倉書店.
- 土屋隆裕 (2010) 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」への適用一, 統計数理, 第 58 巻第 1 号, 25-38.
- 土屋隆裕 (2012) 明治末期における小学生の理想人物調査 -キャリブレーション手法の比較一, 統計数理, 第 60 巻第 1 号, 219-234.
- 土屋隆裕 (2013) 多摩地域住民意識調査 -立川市・小平市郵送調査(2012)-, 統計数理研究所調査研究レポート 108.