

## NFRJ08 のデータ特性： 予備標本・回収率・有配偶率

稲葉 昭 英

第3回全国家族調査(NFRJ08)が2009年1月に実施され、2010年4月より日本家族社会学会内の有志によるデータの共同利用が開始されている。今後、このメンバーによる2次報告書が2011年7月には刊行される予定である。

当初から反復横断調査として計画されていたNFRJにとって、3回目の調査は「趨勢分析が可能になる」という期待を生んでいる。しかし、本当に三つのデータは比較可能なのだろうか？本格的な分析に先立って、まずはNFRJ08のさまざまな特性を把握しておくことが必要だろう。この小稿では、予備標本の使用の影響、年齢別の回収率・有配偶率などを前2回のデータと比較することでこの課題に応えたい。

### I. NFRJ08 データにおける 予備標本の使用

#### 1. 正規標本と予備標本

回収率を算出する前提として、予備標本の影響を確認しておくことが必要とされる。

NFRJ08では、都市規模(18大都市、人口10万以上の市、町村および人口10万未満の市)と都道府県別人口によって層化(108層)を行ったうえで、計画標本9,400をこれらの層に比例配分させ、層別の計画標本を定めている。各層では割り当てられた標本数から、1地点あたり21人を標準として必要な地点数を求め、当該区域の国勢調査区基本単位区から確率比例抽出によって調査対象地点を抽出し、抽出された地点内で必要な標

いなば あきひで：首都大学東京人文社会系

本を抽出するという手続きをとっている(田中, 2009)。

こうして抽出された計画標本は正規標本とよばれる。しかし、これらの標本が死亡や転居などで調査不能であることが実査開始後に判明することがある。こうした場合、NFRJでは余分に抽出しておいた他の標本(予備標本)による代替を認めてきた。NFRJ98では、68~77歳の正規標本が「転居」「死亡」「身体的理由による調査不能」である場合に、予備標本は母集団と同じ性別・年齢階級に属する範囲で選択され、使用された。

NFRJ03では予備標本の使用について方法論的な側面から検討がなされ、1地点につき2票を限度として予備標本を抽出<sup>(1)</sup>、28~77歳の正規標本が「転居」「死亡」「身体的理由による調査不能」である場合に予備標本が使用された。選択された予備標本も不適切であることが判明した場合は、さらに予備標本を使うことはせず、その時点で調査を打ち切ることとした(田中, 2005)。

NFRJ08は、「転居」「死亡」「住所不明」の場合に予備標本の使用を認め、予備標本が不適切である場合には、さらに予備標本の使用を認める、という形をとっている。地点での標本の抽出は年齢や性別による層化をかけていないために、予備標本は28歳から72歳までの男女から無作為に抽出される。予備標本の利用は、抽出台帳の不備などに起因する回収率の低下を改善することになるが、一方で協力的な傾向をもつ人に回答が偏るなどセレクションバイアスを引き起こす可能性が考えられる。

表 1 NFRJ08 における予備標本の使用状況

計画標本	回収 (a)	未回収 (b)	予備標本数 (c)	回収率 (a/(a+b))
正規のみ	4,901	3,996	0	55.1
正規+予備 1	297	192	489	60.7
正規+予備 2	5	7	24	41.7
正規+予備 3	0	2	6	0
合計	5,203	4,197	519	55.4

2. NFRJ08 における予備標本の使用状況

では、NFRJ08 について、正規標本と予備標本の使用状況について概観してみよう。

表 1 のように、計画標本は四つに分類できる。「正規のみ」は、予備標本に代替することなく回収または未回収（拒否など）で終了したもの、「正規+予備 1」は正規標本が予備標本に 1 回代替され終了したもの、「正規+予備 2」は予備標本をさらにもう一度予備標本で代替し終了したもの、「正規+予備 3」はそのまたさらに、予備標本で代替を行って終了したものである。予備標本の使用回数は最も多い場合で 3 回ということになるが、最終的には未回収で終了しており、回収票についてみれば、2 回が最大、大多数は 1 回の予備標本の使用という結果になっている。

使用された予備標本の総数は  $489 + 12 \times 2 + 2 \times 3 = 519$  である。これは複数の予備標本を使用した場合にそれらをすべて含めた数である。重複分を除いて、正規標本のうち予備標本に代替された標本数は  $489 + 12 + 2 = 503$  であり、全体の 5.4% と無視できない大きさである。予備標本 503 のうち、最終的に回収可能であったものは 302 (58.2%) であり、「正規のみ」標本内の回収率 55.1% よりやや高い。この傾向は予備調査票 1 回使用の場合に顕著で、60.7% と「正規のみ」に比して 5 ポイントほど高い値を示している。

3. 予備標本の使用と計画標本

性別年齢階級別にみた予備標本の使用状況（正規標本のうち予備標本に置き換えられた標本数）を表 2 に示す。男女ともに予備標本は最も若い 28-32 歳層で正規標本の 1 割以上を占めている

が、高齢層では 3% 前後と低い傾向が看取できる。男女差はそれほど大きくはないが、男性のほうが概して予備標本の使用率が高い傾向を確認できる。

それでは、予備標本が使用されたことで、当初の計画標本と最終的な計画標本はどの程度異なったものになるのだろうか。表 3 に、性別年齢階級別にみた正規標本と、予備標本を使用した最終標本（複数の予備標本を使用している場合は、最後の標本で代表させる）の度数分布を示した。両者の差異を示す *F-I* の数値からも明らかであるが、予備標本の導入によって男女ともに若年層の標本が減少し（特に男性）、男性の高年齢層および 48 歳以降の女性の標本が増加している。性別では、男性の標本が 61 減少し、女性の標本が同数分増加している。

前節で検討したように、これらの現象は予備標本が繰り返し使用されることでセレクション効果

表 2 性別年齢階級別予備標本の使用状況

年齢	男性	女性	計
28-32	57 (11.8)	50 (10.1)	107 (10.4)
33-37	51 ( 9.2)	31 ( 5.6)	82 ( 7.4)
38-42	40 ( 8.1)	26 ( 5.0)	66 ( 6.5)
43-47	27 ( 5.4)	25 ( 4.9)	52 ( 5.6)
48-52	30 ( 5.9)	12 ( 2.5)	42 ( 4.2)
53-57	29 ( 5.4)	21 ( 4.0)	50 ( 4.7)
58-62	34 ( 5.3)	16 ( 2.4)	50 ( 3.8)
63-67	17 ( 3.3)	11 ( 2.1)	28 ( 2.7)
68-72	12 ( 2.9)	14 ( 3.2)	26 ( 3.0)
合計	297 ( 6.4)	206 ( 4.3)	503 ( 5.4)

(注) 数字は度数 (%), % は当該の人口集団における予備標本の百分率。

表 3 性別年齢階級別にみた正規標本と最終標本の異同

年齢	男性					女性				
	正規標本 (I)		最終標本 (F)		F-I	正規標本 (I)		最終標本 (F)		F-I
	n	%	n	%		n	%	n	%	
28-32	483	10.38	444	9.67	-39	493	10.39	477	9.92	-16
33-37	554	11.91	540	11.76	-14	554	11.67	548	11.40	-6
38-42	497	10.68	483	10.52	-14	522	11.00	527	10.96	5
43-47	502	10.79	499	10.87	-3	515	10.85	517	10.75	2
48-52	510	10.96	501	10.91	-9	489	10.30	509	10.59	20
53-57	534	11.48	524	11.41	-10	526	11.08	538	11.19	12
58-62	646	13.88	648	14.11	2	668	14.07	686	14.27	18
63-67	509	10.94	519	11.30	10	536	11.29	550	11.44	14
68-72	418	8.98	434	9.45	16	444	9.35	456	9.48	12
合計	4,653	100.00	4,592	100.00	-61	4,747	100.00	4,808	100.00	61

が結果として高まったというよりも、最初の予備標本の選択の際に、高齢の男性や中高年女性が多く抽出された結果として生じたと考えたほうがよい。換言すれば、回収が容易であると見込まれる予備標本が調査員によって意図的に抽出された可能性はきわめて高い。正規標本と最終標本の性別年齢別分布の差異はそれほど大きなものではないが、予備標本を使用した結果として、より回答が得られやすいと思われる属性が多く標本に含まれるようになったことは否定できない。

このような結果からすると、予備標本の使用に慎重な統計数理研究所の姿勢(坂元, 2001)はおおむね正しいといえるし、予備標本を使用する際には少なくとも性別は同じ、年齢はできるだけ近い標本を採用するなどのルールを決めておくべきであったのかもしれない。とはいえ、その使用が全体に与えた影響は決定的なものというよりは小さな範囲であり、データを大きく歪めたというほどのものではない。

## II. NFRJ08 データの回収率と未回収の状況

### 1. 性別年齢別回収率

前節で示した最終標本から回収率を計算することができる。この最終標本を用いた性別年齢別回

収率を表 4 に示す。比較のために、NFRJ98, 03 についても示した(NFRJ98 と 03 は 72-77 歳層からもデータを得ているが、この部分は除外している)。なお、三つの調査票ごとに回収率を計算すると、若年者票(28-47 歳)は全体 54.7%, 男性 51.3%, 女性 58.0%, 壮年者票(48-62 歳)は同 56.1%, 54.6%, 57.6%, 高年者票(63-72 歳)は同 55.2%, 55.2%, 55.3%という結果となる。

男性について三つのデータを比較してみると、NFRJ08 がこれまで 2 回の NFRJ とパターンを大きく異にすることが理解できる。NFRJ98, 03 では、回収率は若年層に低く、高齢層で高いというパターンが顕著に示されているが、NFRJ08 ではこの傾向はかなり弱く、年齢階級間の差は小さい。若年層の回収率の低さはこれまで以上に深化しているが、一方で高齢層での回収率が前 2 回に比して大きく低下している。最も回収率が高いのは 58-62 歳、68-72 歳の 56%強であり、NFRJ98 で最も回収率が低かった 28-32 歳層の約 55%とほぼ等しい。68-72 歳層の回収率は、NFRJ98 のそれよりも 20 ポイント以上低下しており、大幅な低下が示されたことになる。また、28-32 歳、38-42 歳層では回収率が 50%を割り込んでいる。

女性についても NFRJ08 の回収率の全般的な

表4 NFRJ98, 03, 08の性別年齢別回収率(%)

年齢	男性			女性		
	NFRJ98	NFRJ03	NFRJ08	NFRJ98	NFRJ03	NFRJ08
28-32	55.0	53.4	48.7	62.8	64.0	55.8
33-37	64.5	55.8	53.7	72.4	68.3	58.4
38-42	62.6	57.8	48.2	69.2	75.2	56.0
43-47	61.9	57.8	54.1	68.8	69.6	61.7
48-52	61.8	57.2	52.1	66.1	61.9	60.3
53-57	61.5	57.3	54.2	70.4	60.0	58.9
58-62	64.0	62.1	56.9	66.5	69.3	54.5
63-67	68.8	68.1	53.8	67.4	65.9	56.0
68-72	80.3	65.8	56.9	72.9	67.7	54.4
合計	63.7	59.3	53.3	68.3	66.6	57.3

低下が著しいことがわかる。過去2回のデータでは高齢層のほか、30代から40代半ばまでの年齢層の回収率が比較的高かったが、NFRJ08ではこれらの層でも大幅に回収率が低下し、ほとんどの年齢層で回収率は55%から60%の数字を示している。ここでも高齢層の回収率の低下は著しく、前2回のデータと比較すると60代以降では10から15ポイント近く回収率が下落している。

以上のように、三つのデータを比較する場合に、回収率に大きな差異があることには十分注意しなくてはならない。三つのデータを比較して何らかの差異が示されたとしても、それが実質的な変化を示しているのか、非標本誤差などに起因する結果なのかはNFRJのみで判断を下すことはできない。他のデータと併用して解釈の妥当性を慎重に判断することが求められる。特に女性と高齢層では大幅に回収率が低下しており、これらの人々に関する比較を行う場合、この点には留意する必要がある。

## 2. 欠票理由

回収率の低下はなぜ起きたのだろうか。最終標本において、未回収となった標本の欠票理由を検討してみると、「拒否」が圧倒的に多く(70.2%)、次いで「一時不在」(17.1%)で、両者をあわせて8割以上、というパターンは2回の調査でそれほど変わりはない[NFRJ03ではそれぞれ67.2%、

17.3%(田中, 2005)]。NFRJ08で3ポイントほど拒否が増加しているが、この数字からは極端な変化があったとは言い切れない。

NFRJ08において「拒否」「一時不在」の二つの理由の占める比率を性別年齢階級別に比較してみると、全般的には男女ともパターンは似ており、それほど大きな差があるとはいえない。若年では「一時不在」が3割近くまで達するが、年齢の上昇とともにこの比率は低下し、高齢層では「拒否」の比率が男女ともに7割以上に達する。「拒否」の傾向は、「一時不在」がやや低いことと連動するのか、女性に高い。

こうしたパターンをみる限り、NFRJ08で発生した高齢層および女性の回収率の低下は、基本的には「拒否」によって生じていると考えるほかないようだ。社会調査の回収率の長期的な低下が「拒否」の増加に起因することは、すでにSynodinos and Yamada (1999)によって指摘されている。2000年以降の社会調査の回収率の低下には、「振り込め詐欺」などの犯罪の増加が関与していることも指摘されている(Inaba, 2007)。従来は女性と高齢者が社会調査の回収率を保つ有力な存在であったが、これらの犯罪は女性と高齢者をターゲットとしたため、多くの被害が報道されるなかでこれらの層にも社会調査への拒否傾向が高まった可能性は否定できない。実際に、「社会意識

に関する世論調査」(総理府)や「外交に関する調査」(外務省)についても、2005年以降に女性と高齢者の回収率の大幅な低下が示されている(Inaba, 2007)。もしこの仮定が正しい場合には、高齢層の回答者と回答拒否者の間に意識や態度面で何らかの差異が存在することが考えられる。

### III. 三つのデータの有配偶率

NFRJ08データの特性は官庁統計などとの比較によってある程度検討可能であるが、この作業は永井論文に譲り、ここでは過去2回のNFRJデータと婚姻上の地位に関する比較を試みる。家族研究にとっては有配偶率のもつ意味はきわめて大きいからである。

表5に、男女別に有配偶率を示した。参考までに、2005年の国勢調査の結果も示してある。まず、三つのNFRJデータを比較すると、男女ともに有配偶率はNFRJ98・03が比較的類似していたのに対して、NFRJ08では大きな低下を示している。28-39歳についてみると、男性では6.2ポイント、女性では8.6ポイントと大きな低下を示し、40代でもこの傾向は続く。一方で女性の60-72歳については過去よりも有配偶率が高いという逆転現象が生じている。

過去2回と比較してみると、男性はどの年齢層でも有配偶率が低下している。とはいうものの、2005年の国勢調査と比較すると、それでもどの年齢層でも5から8ポイントほど高い。2005年

国勢調査と最も調査時点に近いのは2004年1月に実査が行われたNFRJ03であるが、NFRJ03と国勢調査との差異はどの年齢層でも大きく、最大で12.8ポイントと、有配偶者の比率がNFRJ03において過大であることがわかる。

以上の傾向は女性にも同様に示される。NFRJ08の有配偶率は、過去2回のNFRJと比較して大きく低下しているが、2005年国勢調査と比較するとそれでも高い。有配偶者が過大である傾向は男性と同様にNFRJ03で顕著であり、格差は最大で13.2ポイントにも達する。

このように、NFRJデータは一貫して有配偶者が過大に含まれていることになるが、三つのデータを比較すると、この傾向はむしろNFRJ03に顕著であるように思われる。NFRJ08は過去よりも有配偶者の比率が低下しているが、おそらく母集団との差異はNFRJ03よりは小さいと推測できる。

NFRJデータには有配偶者が過大に含まれることはこれまでも指摘されてきた(稲葉, 2004)。それでも、家族研究は有配偶者や子どものいるものを対象に分析をすることが多かったために、有配偶率の変化は多くの分析であまり深刻な問題とはなっていない。しかし、そうした対象に限定せずに分析を行うことが多い、態度や意見に関する項目を分析する際には、この点を適切に統制する必要は大きい。同時にNFRJ03の過大な有配偶率は、あらためてその原因と影響を慎重に検討する必要を物語っている。

### IV. おわりに

以上、NFRJ08データについて予備標本を使用したことの影響と、回収率、および過去のNFRJデータと比較したうえでの有配偶率について略述した。

予備標本を使用した結果として最終的な標本には若年男性が減少し、中高年女性が増加した。しかし、最終標本の分布をみる限りは、予備標本の使用が標本構成を大きく変えるということはない。

表5 NFRJデータの有配偶率(%)

年齢	NFRJ98	NFRJ03	NFRJ08	2005 国調
男性				
28-39	67.0	65.5	60.8	52.9
40-49	85.5	86.3	80.0	72.6
50-59	89.7	86.8	84.1	79.1
60-72	93.7	90.9	89.3	84.8
女性				
28-39	76.2	77.2	67.6	64.0
40-49	87.8	86.1	82.9	79.0
50-59	86.9	86.4	83.3	80.1
60-72	73.1	73.8	76.2	70.7

かった。予備標本の使用には課題を残すが、その影響はとりあえず深刻なものとは考えなくともよいようだ。

回収率に関しては、NFRJ08 では高齢者と女性の回収率の低下が著しかった。高齢層と女性について、過去 2 回のデータと比較して分析を行う場合には、NFRJ08 の結果が非標本誤差の影響で生じている可能性も含めて慎重な分析が必要だろう。

有配偶率については調査ごとに低下しているが、それでも国勢調査などの結果よりは高い。分析の結果はむしろ、NFRJ03 において有配偶率が母集団と比較して著しく高いことが明らかとなった。

こうした結果からすると、1 変数の分布を三つのデータで単純に比較することはかなり危険であることがわかる。配偶者の有無や年齢、性別の統制はもちろんのこと、他のデータを併用しつつ NFRJ で示された結果が他のデータでも支持されるかどうかを検討するような、慎重な姿勢が望まれる。

#### 【付 記】

NFRJ08 データは全国家族調査委員会の許可を得て使用した。なお、本稿中の予備標本の使用状況に関する

分析は、田中重人氏（東北大学）によるデータの整理によって可能になった。田中氏にお礼申し上げる。

#### 【注】

- (1) 1 地点に 2 票までとしたのは、予備標本を多く抽出しておく、現場の調査員が「調査拒否」などのケースに予備標本を使用しかねない、という危惧からの措置である。

#### 【文 献】

- 稲葉昭英, 2004, 「NFRJ98 の調査設計とデータ特性」  
渡辺秀樹, 稲葉昭英, 嶋崎尚子編『現代家族の構造と変容』東京大学出版会, 15-24.
- Inaba A., 2007, "Problems relating to declining response rates to social survey research in Japan: Trends after 2000," *International Journal of Japanese Sociology*, 16: 10-22.
- 坂元慶行, 2001, 「日本人の国民性調査: 社会調査研究の最前線」『理論と方法』16: 75-88.
- Synodinos N. E., Yamada S., 1999, "Response rate trends in Japanese surveys," *International Journal of Public Research*, 12: 48-72.
- 田中重人, 2005, 「サンプリングとデータの基本特性」  
日本家族社会学会全国家族調査委員会編『第 2 回家族についての全国調査 (NFRJ03) 第一次報告書』, 23-34.
- 田中重人, 2009, 「NFRJ08 標本抽出と調査実施」『家族社会学研究』21(2): 208-213.