

NFRJ08 における複数調査票の作り方

保 田 時 男

I. 問題の所在

1. なぜ複数の調査票を用いる必要があるのか

全国家族調査 (National Family Research of Japan; 以下, NFRJ) は, 日本家族社会学会が学会企画として長年取り組んでいる全国規模の代表サンプルによる継続調査である。これまで2回の調査 (NFRJ98, NFRJ03) がそれぞれ1999年1月, 2004年1月に実施され, 現在, 2009年1月の実施をめざして第3回調査 (NFRJ08) の準備が進んでいる⁽¹⁾。NFRJ08の準備作業は, 学会員有志によるNFRJ08実行委員会が担っており, 2006年夏から具体的な議論が始まっている。

筆者は, 実行委員の1人として, 調査全体のデザイン (調査対象や調査項目の枠組み) を検討するデザイン班の活動を行っている。本稿は, 「NFRJ08において, 複数の調査票をどのように設定するか」という問題に対する取り組みの過程を紹介するものである。このような調査の準備作業での取り組みについて, 当事者以外の目に触れることは少ないが, 公開データとして共有されることが前提のNFRJにおいては, その過程が明らかにされることには一定の意義があるであろう。

NFRJは1回の計画サンプル・サイズが10,000人に及ぶ大規模調査であり, 28~77歳という幅広い年齢層の男女を全国から抽出している。広範囲のサンプルをもつ大規模調査をデザインするには, 想定される回答者が多様性に富むため, 尋ねるべき調査項目を全体として統一しにくいことが一つの問題となる。回答者のライフ

ステージによって調査すべき事柄が大きく変化する家族調査においては, このことは特に問題になる。通常, この問題には枝分かれ質問で一部の調査項目について回答者を限定することで対応するが, 枝分かれが多すぎると調査票が複雑になる。NFRJは留置調査であるため, 回答者が混乱するような複雑な枝分かれは避けなければならない。

そこで考えられる方策が, あらかじめ多様な回答者にみあった複数の調査票を用意しておき, 回答者の属性に合わせて適切な調査票を割り当てることである。前回調査 (NFRJ03) では, サンプルを47歳以下と48歳以上で2つに分割し, それぞれに若年調査票と中高年調査票を割り当てていた。NFRJ08において同様に2つの調査票を用いるべきかどうか, あるいは別の方法を採用すべきか, その決定は調査デザイン全体に影響するので, 実行委員会の中でも早期に議論が求められた。

2. NFRJ03で用いられた複数の調査票

前回のNFRJ03では, 若年調査票と中高年調査票の2種類が用いられたが, この2分割は, あまり機能的な役割を果たしていたとは思えない。なぜならば, NFRJ03の2つの調査票はその内容がほとんど共通で, 異なる個所はごく一部に限られていたからである。

このように2つの調査票が似通っているのは, そもそも2つの調査票を異なる内容にしようとする意志が弱かったとも考えられるが, 内容の違いを設けることが難しかったためと考えられる。47歳以下と48歳以上の2分割の場合, 47・48歳をまたぐ年齢層が対象となる研究テーマを扱うための調査項目は, 結局, 両方の調査票に組み込ま

やすだ ときお: 大阪商業大学総合経営学部



2. 結果

図1がクラスター分析の結果にもとづいて、86本の論文を並べ替えたものである。1つの行が1つの論文を表しており、中央に論文番号を記している（例えば「03-2-13」はNFRJ03報告書No.2の13本目の論文である）。中心から左が男性、右が女性を示しており、網がかかっている部分が分析対象になっている年齢である。図の右側には、クラスターの構造を示すデンドログラムを記している。

この図を上から順に観察していくと、次のよう

なことがわかる。上から39本目までの論文は、男女ともすべての年齢の回答者を分析対象としている（全年齢パターン）。40～43本目も若干高齢者が削られているものの、ほぼ全年齢が分析対象である。44～46本目の論文は、若齢層を分析対象外とし、およそ40歳以上を分析対象にしている（壮齢～高齢パターン）。47～55本目の論文は女性のみを対象にし、56～60本目の論文は男性のみを対象にしている。ここでは、男女にかかわらず対象年齢のパターンに注目しているので、いったん読み飛ばそう。60歳以上の男性のみを

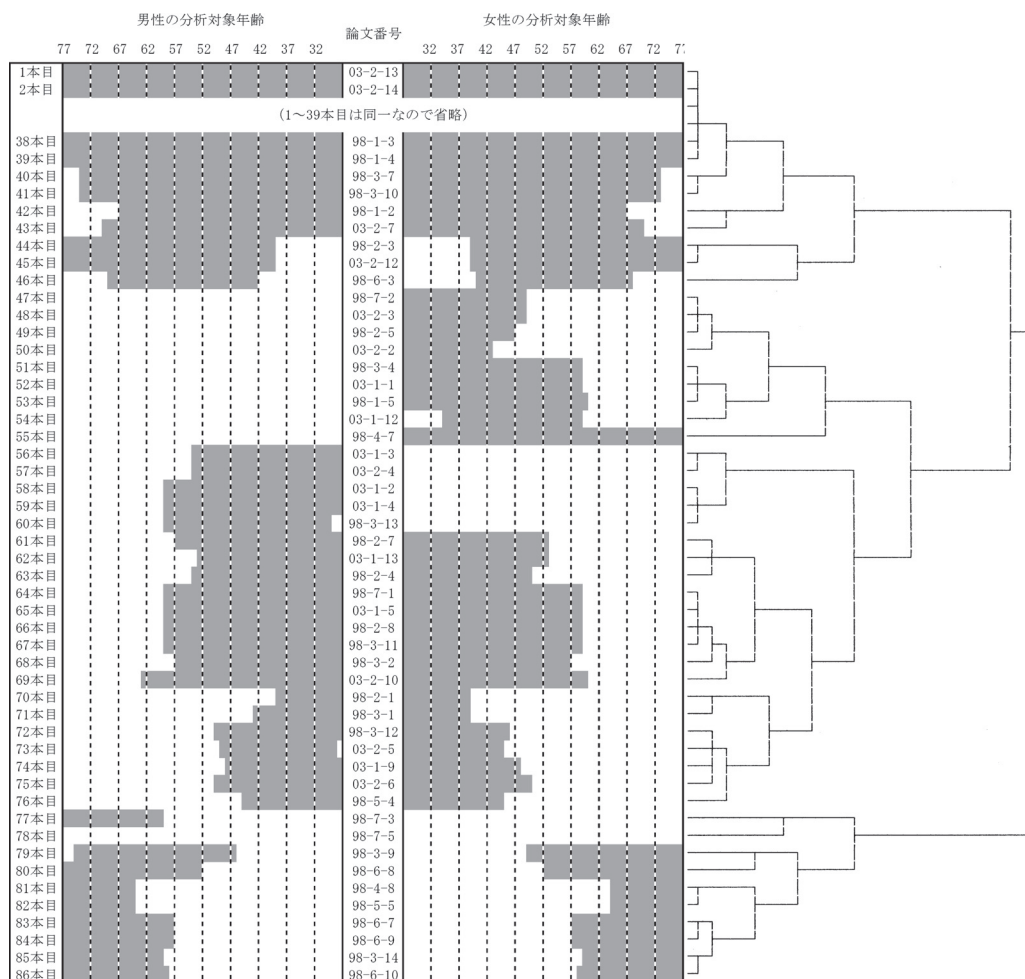


図1 報告書論文のクラスター分析の結果



対象とする77本目の論文や、分析対象がない（データ分析を含まない）論文である78本目の論文も同様に読み飛ばす。61～69本目の論文は、おおまかに60歳程度までを分析対象としている（若齢～壮齢パターン）。70～76本目の論文はせいぜい40代程度までの比較的若い層のみを分析対象としている（若齢パターン）。79～86本目の論文はやや開始年齢にぶれが大きいもののおおまかに高齢者のみを分析対象にしているといえる（高齢パターン）。

以上のように観察すると、分析対象とする年齢層の限定の仕方には、大まかに、全年齢パターン、若齢パターン、若齢～壮齢パターン、壮齢～高齢パターン、高齢パターンの5つがあることがわかる。男性のみ、女性のみを分析対象としている論文も、47～50本目は若齢パターン、51～54本目は若齢～壮齢パターン、55本目は全年齢パターン、56～60本目は若齢～壮齢パターン、77本目は高齢パターンが当てはまるであろう。

ここで注目すべきことは、一連の報告書論文は分析の目的も方法も多様であるにもかかわらず、分析対象とする年齢層の区切り方にはかなりの類似パターンが認められるということである。しかも、直接に年齢層で分析対象を限定していない（婚姻状態や子どもの年齢で分析対象を限定している）論文を含めて、このパターンは共有されている。このクラスター分析の結果は、年齢層による調査票の分け方について、多様な分野の家族研究者の間である程度共通の認識が得られそうなことを示唆している。

III. NFRJ08の複数調査票デザイン

1. 若齢・壮齢・高齢調査票

クラスター分析によって、年齢層による分析対象の限定には5つのパターンがあることが明らかにされた。全年齢パターンを除けば4つであるが、では4つの調査票を用意すればよいのかというと、それほど単純ではない。それぞれのパターンの中に重なりがあるためである。例えば、若齢

の回答者には、若齢パターンと、若齢～壮齢パターンの両方が当てはまる。

具体的に複数調査票の設定を考える際には、シカゴ大学NORCのGeneral Social Survey（以下GSS）で採用されている調査項目のローテーション・デザインが参考になった。GSSは多くの調査項目を継続的に調査しているので、中心的な調査項目だけに絞っても、毎回すべての回答者にすべての調査項目を尋ねることは、回答者に大きな負担となる。しかし、今回はAグループの調査項目、次回はBグループ、その次はまたAグループ、……と単純にローテーションで回してしまうと、Aグループの調査項目とBグループの調査項目の関係性を調べる分析が永遠にできないことになってしまう。そこで、GSSでは1988年調査以降、次のような方法を採用している。まず、調査項目をA、B、Cの3つのグループに分け、また回答者もランダムに3つのサブサンプル(ballot)に振り分ける。そして、1つ目のサブサンプルにはA+Cグループの調査項目に回答してもらい、2つ目のサブサンプルにはA+Bグループ、3つ目のサブサンプルにはB+Cグループの調査項目に回答してもらう。これを毎回の調査で繰り返せば、それぞれの調査項目の回答者は全体の3分の2に限られるが、すべての調査項目について毎回調査し、かつどの組み合わせの調査項目についても（3分の1の回答者のデータを利用して）分析を行うことができる(Smith, 1988)。

このGSSのローテーション・デザインは、調査項目のグループと回答者のサブサンプルを1対1に対応させるのではなく、複数の調査項目グループを組み合わせることで回答者のサブサンプルと対応させることに、大きなメリットがあることを示唆している。NFRJ08の複数調査票のデザインにおいても、このことを前提に、複数の調査項目グループを組み合わせるという考え方で、クラスター分析の結果に対応するデザインを考慮した。

結論として、筆者はNFRJ08では図2のように複数調査票を設定することを、実行委員会に提

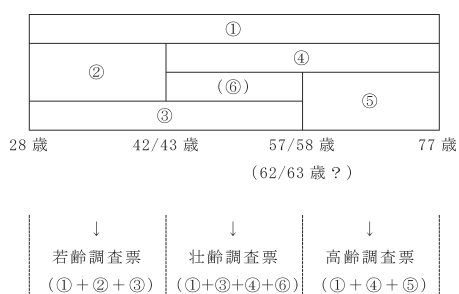


図3 NFRJ08 で用いる複数調査票 (改訂案)

議論となった。報告書論文のクラスター分析を見る限り、壮齢のみを対象とする分析の需要は多くないので、壮齢のみを対象とする調査項目を設定する必要性は強くない。しかし、図3の⑥のように壮齢のみの調査項目を設定したとしても、この複数調査票のデザイン上問題にはならないので、壮齢のみの調査項目の可能性も残しておいてよいであろうという結論に至った。

最後に、いわゆるキャリーオーバー効果など、調査票の構成に依存する文脈効果(context effects)の懸念に注意を払う必要がある。例えば、グループ③の調査項目は、若齢調査票にも壮齢調査票にも含まれるが、2つの調査票では他に含まれる調査項目が異なるので、前後の調査項目の並び順などの影響を受けて、回答傾向に異なるバイアスがかかる恐れがある。GSSのローテーション・デザインにおいては、このような文脈効果によるバイアスが発生していないかどうか事後的な分析で検討しており、ほとんどの調査項目について皆無もしくは非常に小さい文脈効果しか起こっていないことを確認している(Smith, 1991)。しかし、同じことがNFRJにも期待できるかはわからない。さらに注意すべきことは、GSSのローテーション・デザインのようにランダムに回答者のサブサンプルを作成する場合と違って、NFRJ08における回答者のサブサンプルは年齢層によって区切られるものなので、2つの調査票で回答の傾向が異なったときに、年齢層の違いによるものなのか、文脈効果のバイアスによるもの

なのか事後的に区別をすることが難しいということである。このような点を考えると、文脈効果の影響を受けやすい意識・態度項目を、複数の調査票にまたがるグループ③や④に割り振る際には、前後の調査項目の配置に特に注意を払うべきであろう。

本稿では、NFRJ08の調査デザインについて、実行委員会での取り組みの一端を紹介した。NFRJはデータ公開を前提とするので、その分析者は必ずしも調査設計やデータ収集にかかわっていない。しかし、実際に公開データの分析を行う際には、その調査過程も含めて、NFRJという調査に慣れ親しむ必要がある。一般に、自らが実施していない調査に慣れ親しむには膨大な時間がかかり、そのことは公開データの利用をためらわせる大きな理由となる(Hofferth, 2005)。調査準備の過程を紹介することが、NFRJのデータ利用者にくらかの情報を示し、スムーズな分析の道を開くことになれば幸いである。

【注】

- (1) 末尾の2桁の数字はサンプリング時点の年度を表している。NFRJは実査前年の12月31日時点での年齢にもとづいてサンプリングを行っているので、2009年実施の第3回調査は、NFRJ08となる。12月31日時点での年齢にもとづくサンプリングには、生まれ年と年齢が1対1に対応し、分析上扱いやすいという利点がある。
- (2) 例えば、60歳以上の男性を分析対象とする論文の場合、……59歳男性=0, 59歳女性=0, 60歳男性=1, 60歳女性=0……といった具合である。
- (3) 論文中に示されているとおりに分析対象を限定しても、ケース数が論文の記述とは厳密には合致しないことがいくつかあった。これは、論文中の記述漏れ、あるいは筆者の見落としのためと考えられるが、本稿の目的を果たすためには大枠で間違いがなければ問題は無いので、小さな齟齬は追及していない。

【文献】

Hofferth S.L., 2005, "Secondary data analysis in

研究動向：NFRJ08における複数調査票の作り方



family research," *Journal of Marriage and Family*,
67(4): 891-907.
Smith T. W., 1988, "Rotation designs of the GSS,"
GSS Methodological Report, No. 52.
Smith T. W., 1991, "Context effects in the general

social survey," In Biemer P.P., Groves R.M.,
Lyberg L.E., Mathiowetz N.A., Sudman S. (ed.),
Measurement errors in surveys, John Wiley & Sons,
Inc., 57-72.

