

世帯統計の世帯員構成を簡単に判別できるシステムの開発と利用法

(公財) 統計情報研究開発センター

周防節雄

「官民オープンデータ利活用の動向及び人材育成の取組」研究集会

独立行政法人統計センター (オンライン研究会)

2022年11月24日

要約

昨年と一昨年の本研究集会では、世帯を調査対象とする公的マイクロデータの世帯構成を的確に表現できる「家族構成変数」の開発と利用例について、研究報告を行った。今回は、一般ユーザーがその家族構成変数を使って、どんなに複雑な世帯構成でも、SAS で簡単に抽出できるアルゴリズムを開発したので紹介する。更に、国勢調査の匿名データに適用して、いくつかの抽出結果例を提示することで、その有用性を示す¹。

1. はじめに

「家族構成変数」については、昨年と一昨年の本研究会で研究報告をしたので、この変数の概略についてはそちらを参照されたい。

我々のグループでは、8050 問題、少子高齢化、人口減少、ひとり親世帯、貧困世帯、経済格差など、現在、早急に対応が迫られている課題に関心を持っている。これらの諸問題の対策には、各世帯の家族構成が重要なキーとなる。

世帯を調査客体とする統計調査は数多くあり、(独)統計センターで作成される調査としては、国勢調査(以後、「国調」)、労働力調査、全国消費実態調査、就業構造基本調査、社会生活基本調査等がある。これらの調査では、各世帯員の4つの個人属性、①続柄、②婚姻状況、③性別、④年齢が調査されており、各世帯の世帯員構成が特定できる。続柄の表現には調査によって若干の相違はあるが、ユーザー側で統一できると考えている。

各統計調査には、独自に定義された世帯分類変数が含まれている。例えば、国調には「世帯の家族類型」と称する変数が昭和 45 年調査から作成・提供されてきた。そこで分類されている家族類型のひとつに「夫婦と子供から成る世帯」がある。その中味を見ると、大きく分けて、①「世帯主夫婦と子供から成る世帯」、②「世帯主と両親から成る世帯」、③「世帯主と両親、および世帯主の兄弟姉妹から成る世帯」の3種類の世帯構成が混在しているが、①と②は明らかに社会生活上異なる性格の世帯である。「夫婦と子供から成る世帯」といきなり言われると、殆どの方は①だけを頭に浮かべて、分析することが予想され、分析結果の解釈に戸惑うことになる。ただ、誤解を避けるために一言付け加えると、この国調の「世帯

¹ 本論文は、2022年11月24日に開催されたオンライン研究会での報告用パワーポイントを基にして、解説を加えて作成され、web にアップロードされた。そのため、モノクロではなく、カラー表示をそのまま採用した。

「家族類型」の変数を作成するにはかなり複雑なアルゴリズムが必要であることは、想像に難くないが、エンドユーザに提供する際は、丁寧な解説が必要だと思う。

先に挙げたその他の統計調査でも、それぞれ「独自の」世帯分類変数が用意されていると思われるが、もしそうであるなら、異なる調査間で世帯の行動を比較する際には、何らかの調整が必要となり、これは極めて複雑な処理を伴うことになる。我々が発想した「家族構成変数」は、異なる統計調査間でも、同じ定義の変数で各世帯を分類することが可能になるので、世帯構成毎の行動パターンを容易に、かつ、多面的に比較することが可能になる。

2. 家族構成データベースの作成

(独)統計センターから提供された「国調・匿名データ」を、そのままの構造で分析用 SAS データベースに変換後、世帯情報と世帯員情報の二つのデータセットに分離した(図1)。

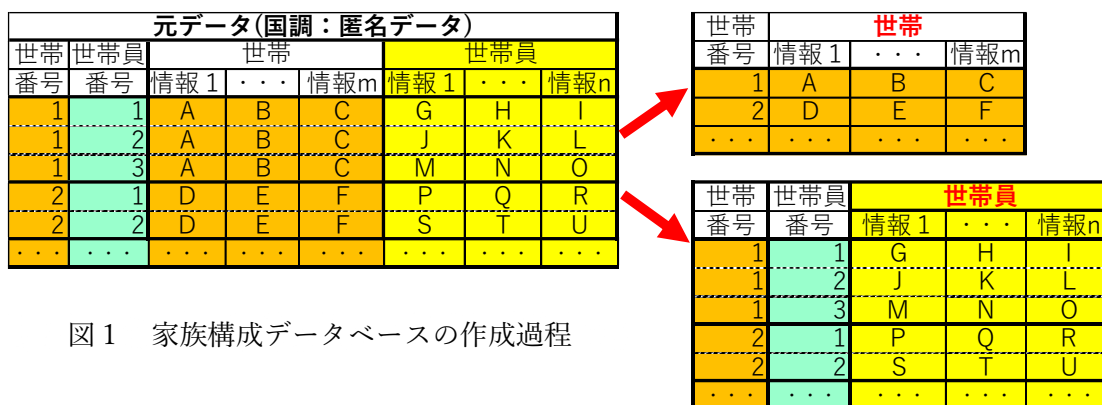


図1 家族構成データベースの作成過程

世帯員情報データセットから、更に、**世帯員続柄データセット**を作成した。その構造例を図2に示す。このデータセットには二種類の**家族構成変数**が含まれている。

- ① **全体版**(変数名: kanjiPat) : 世帯内の全世帯員の続柄を出現順に左から羅列した
- ② **短縮版**(変数名: shortPat) : 同じ続柄は一つにまとめた

この変数に使用される漢字の意味は次節で解説する。

変数	家族構成変数		世帯ID番号 HHidNo	世帯員数 HHmNo	個人属性												世帯主年齢階層 HeadAge	世帯の種類 HHkind	世帯の家族類型 FamilyType	3世代世帯 gene3				
	全体版	短縮版			続柄				性別				年齢階層								婚姻状況			
	kanjiPat	shortPat			tsuzuki				sex				age								marry			
1	夫妻孫	夫妻孫	1	3	1	2	7	1	2	2	15	15	5	2	2	1	15	1	4	0				
2	父母仲伴	父母伴	2	4	1	2	3	3	1	2	1	1	7	7	2	1	2	2	1	1	2	0		
3	嬢	嬢	3	10				2			18			3			18	2						
4	単	単	4	1	1			1			6			1			6	1	6	0				
5	夫妻	夫妻	10	2	1	2		1	2		16	16		2	2		16	1	1	0				

数字1~8は、家族構成変数 kanjiPat の漢字の左からの並びに対応している。 原データから付与した。

【注】3番目の世帯の世帯員数が「・」(欠損値)は、一般世帯の「対象外」を意味する。年齢階層が「18:85歳以上」なので、施設・病院等に入居と推測できる。年齢階層は、国調・匿名データでは5歳刻み。今回、更に、各世帯員の年齢実数をその区間の中間値として、変数xage1~xage8を追加した。

図2 世帯員続柄データセットの構造例

3. 家族構成変数に使用される漢字一覧

国調・調査票情報(2000年以降)で使用される12種類の続柄情報と、性別、婚姻状況、年齢(一歳刻み)を組み合わせて、各世帯員の家族関係を33個の漢字で表現した(図3)。

続柄		(国勢調査・調査票情報用)																																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33				
1 世帯主または代表者	世帯種類	単独世帯				二人以上の世帯																																
2 世帯主の配偶者	世帯主	世帯主				世帯主												世帯員																				
3 子	世帯員	世帯主				世帯主												世帯員																				
4 子の配偶者	続柄コード	1				1				1/2		1/2		3			4		5		6		8		7		9		10		11		12					
5 世帯主の父母	世帯員変数値	単	宿	嬬	嬬	男	女	亭	奥	父	母	夫	妻	倅	倅	娘	好	嫁	婿	爺	袋	舅	姑	翁	婆	孫	兄	弟	姉	妹	威	雇	他	媒				
6 世帯主の配偶者の父母	性別	M	F	F	F	M	F	M	F	M	F	M	F	M	F	F	F	F	M	M	F	M	F	M	F	F	M	M	F	F	M	F	M	F	F			
7 孫	世帯主から見た続柄	单身	女(六)	女(十)	女(十)	女(十)	女(十)	女(十)	女(十)	子・同居	子・同居	世帯主・同居	世帯主・同居	息子・同居	息子・同居	娘・同居	娘・同居	息子の妻	息子の妻	父	母	配偶者の父	配偶者の母	祖父	祖母	孫	兄弟姉妹	兄弟姉妹	世帯主との年齢比較で判別	×	×	×	×					
8 祖父母	婚姻状況	非	既	既	既	非	既	既	不	既	既	既	既	既	既	既	既	既	既	既	既	既	既	既	不	不	不	不	不	不	不	不	不	不				
9 兄弟姉妹	年齢	年齢：一歳刻み																																				
10 他の親族	注	非婚=未婚・死別・離別 性別：×=不問												国勢調査・匿名データでは、威・雇・他・媒は、10に一括りになっているので「威」と表示される。																								
11 住み込みの雇人	匿名データ	匿名データ 10																																				
12 その他																																						

図3 家族構成変数に使用される漢字一覧

この33個の漢字を使用したシステムは、以前、国調・調査票情報を利用した際に開発した。本稿の分析で利用している国調・匿名データでは、続柄コードの10~12は一括りに「10」にまとめられている。つまり、本稿で使用する続柄コードの「10」は、国調・調査票情報の「10~12」を意味しているが、実態は、殆どが「12」の「他」と「媒」と見なしてよいと考える。これら33個の漢字の世帯内での階層構造を図4に示す。

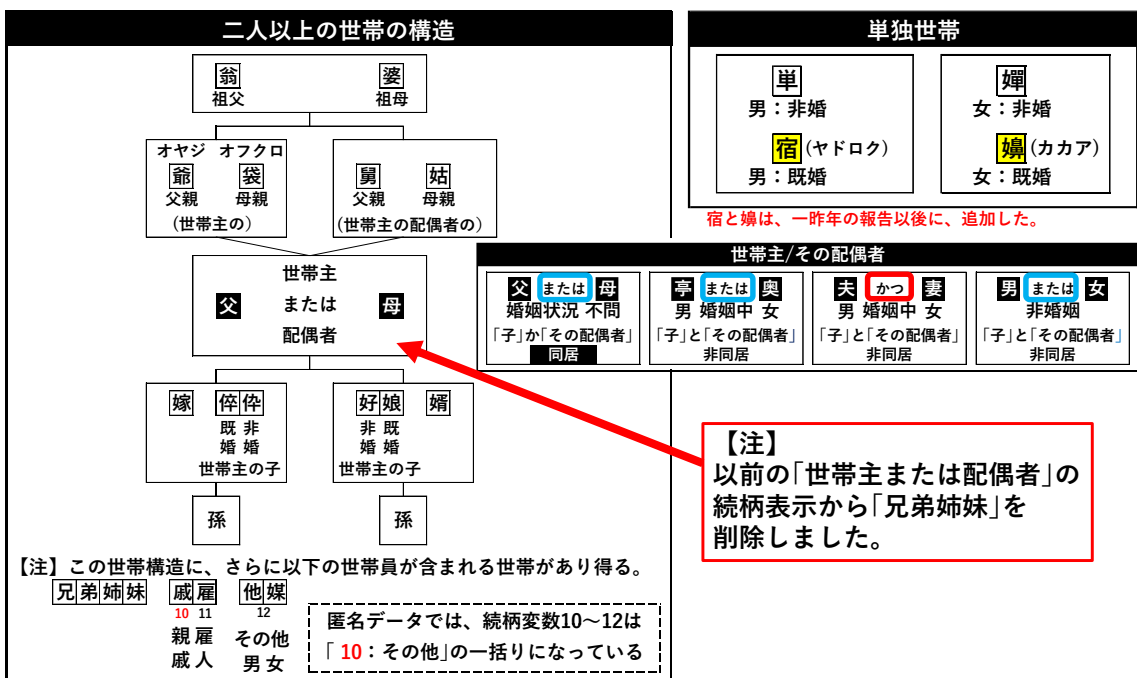


図4 33個の漢字の世帯内での階層構造

4. 国調・匿名データで提供される変数「世帯の家族類型」と家族構成変数の対応関係

我々が開発した家族構成変数の意味を明確に理解する目的で、提供された国調・匿名データに組み込まれている変数「世帯の家族類型」と家族構成変数(短縮版)のクロス表を作成した(表1)。

順位	家族構成変数 (短縮版)	2: 夫婦と子供 から成る世帯
①	父母伴	54,578
②	父母好	47,131
③	父母伴好	45,122
④	男爺袋	509
⑤	父母伴	314
⑥	父母娘	223
⑦	女爺袋	74
⑧	男爺袋妹	50
⑨	父母伴戚	49
⑩	父母好戚	47
11位～36位省略		
①位～⑩位の合計		148,452

一般世帯 合計 (462,355件)
欠損値「施設等の世帯」 (19,649件)

国調・匿名データ (2000年)

1:夫婦のみの世帯
2:夫婦と子供から成る世帯
3:男親又は女親と子供から成る世帯
4:その他の親族世帯
5:非親族世帯
6:単独世帯

世帯の家族類型

表1 変数「世帯の家族類型」と
家族構成変数(短縮版)のクロス表

表1の順位①～⑧だけを詳細に見ると、更に、以下のように分類できる。

- (1) 順位①④⑤：両親と息子（但し、④は、息子が世帯主）
- (2) 順位②⑥⑦：両親と娘（但し、⑦は、娘が世帯主）
- (3) 順位③⑧：両親と息子と娘（但し、⑧は、息子が世帯主で、その妹も同居）

従って、①～⑧は確かに「夫婦と子供から成る世帯」ではあるが、世帯主が親か子で社会・経済活動は大いに異なることは明らかなので、提供されたデータにある変数「世帯の家族類型」を使って得た種々の分析結果はそのまま解釈するには無理がある。更に、順位①と⑤のように、同居する子が独身か既婚かによっても世帯の行動様式は大いに異なるはずである。

5. 家族構成変数による2000年から2005年間の世帯構造の変遷

国調・匿名データの2000年と2005年から作成した家族構成変数(短縮版)ShortPatを使って、この両年間の世帯構造の変遷を分析した。両年のShortPatの出現頻度を算出した後、2005年の頻度表を降順にソートし、その頻度順に2000年のShortPat(同じ続柄の世帯員が複数同居している場合は、一つに圧縮)を対応させた結果表が、表2-1である。

上位11位までは、2005年の順位9位・10位が2000年の順位と入れ替わっている他は、順位不動だったが、増減数に特徴が見られた。

- (1) 1位「夫妻」だけの世帯：8,267世帯の増加
- (2) 2位「単(独身男性：単独世帯)」：7,299世帯増加
- (3) 3位「嬪(独身女性：単独世帯)」：10,392世帯の増加
- (4) 4位～5位「両親と独身の息子又は娘」：ほぼ不変
- (5) 6位「両親と独身の息子・娘」：3,042世帯の減少
- (6) 7～8位「母親と独身の子」：2,925世帯の増加
- (7) 9位「既婚男性の単独世帯」：微増
- (8) 10位「両親と息子夫婦と孫」の三世帯世帯：1,312世帯の減少
- (9) 11位「母(世帯主)と独身の息子・娘」：759世帯の増加

表2-1の「増減数」を見ると、2005年に減少した世帯形態は、27位、30位、31位以外は全て両親が世帯主夫婦の世帯であった。特に、順位6位の「父母倅好」(両親と**独身**²の息子と娘が同居する世帯)が3,042世帯の大幅減になった。

表2-1 国調・匿名データの2000年から2005年間の世帯構造の変遷：全世帯

2000年		2005年					
%	世帯数	shortPat	世帯数	増減数	%	累積%	順位
18.3	87,984	夫妻	96,251	8,267	18.9	18.9	1
14.8	71,206	単：独身男単独世帯	78,505	7,299	15.5	34.4	2
13.9	66,788	嬢：独身女単独世帯	77,180	10,392	15.2	49.6	3
11.3	54,578	父母倅	54,707	129	10.8	60.4	4
9.8	47,131	父母好	46,952	-179	9.2	69.6	5
9.4	45,122	父母倅好	42,080	-3,042	8.3	77.9	6
2.0	9,746	母倅	11,474	1,728	2.3	80.2	7
1.9	8,921	母好	10,118	1,197	2.0	82.2	8
1.6	7,651	宿：既婚男単独世帯	8,124	473	1.6	83.8	9
1.7	8,102	父母倅嫁孫	6,790	-1,312	1.3	85.1	10
1.0	4,895	母倅好	5,654	759	1.1	86.2	11
0.9	4,254	夫妻袋	4,418	164	0.9	87.1	12
0.8	3,711	男袋	4,075	364	0.8	87.9	13
0.9	4,414	父母倅袋	3,884	-530	0.8	88.7	14
0.6	3,046	嬢：既婚女単独世帯	3,679	633	0.7	89.4	15
0.9	4,449	父母倅好袋	3,471	-978	0.7	90.1	16
0.8	3,763	父母好袋	3,305	-458	0.7	90.8	17
0.4	2,106	父倅	2,494	388	0.5	91.3	18
0.3	1,296	男戚	1,932	636	0.4	91.7	19
0.3	1,360	父好	1,512	152	0.3	92.0	20
0.3	1,297	夫妻姑	1,465	168	0.3	92.3	21
0.2	1,081	父母好孫	1,425	344	0.3	92.6	22
0.3	1,342	父母倅嫁	1,339	-3	0.3	92.9	23
0.3	1,439	父母娘婿孫	1,227	-212	0.2	93.1	24
0.2	1,100	女袋	1,209	109	0.2	93.3	25
0.2	1,141	父母倅姑	1,179	38	0.2	93.5	26
0.3	1,218	父倅嫁孫	1,139	-79	0.2	93.7	27
0.3	1,471	父母倅好爺袋	1,065	-406	0.2	93.9	28
0.2	990	父母好姑	1,017	27	0.2	94.1	29
0.2	1,035	女妹	976	-59	0.2	94.3	30
0.2	1,010	母倅嫁孫	956	-54	0.2	94.5	31
0.1	616	女戚	908	292	0.2	94.7	32
中略							
0.0	5未満	奥兄弟	0	-1			1,491

² 現在の我々のシステムでは、世帯主の「子」のうち、「倅」「好」は「現在婚姻(事実婚を含む)中でない」子を意味する。従って、本論文で言及する「独身の子供」とは、世帯主の子で「現在独身者」を意味し、未婚、離別、死別の「子」が該当する。いずれ、今、二種類ある「家族構成変数」の全体版と短縮版に加え、**第三の家族構成変数**として、未婚か結婚歴有りを区別できる新しい続柄用の漢字を、男女一文字ずつ新規導入して、近々、バージョンアップする予定である。

表2-2では、表2-1から「両親と独身の子」の世帯、母親が世帯主の母子世帯、独身の子が世帯主の母子世帯だけを抽出した。

表2-2 国調・匿名データ 2000年から2005年の世帯構造の変遷：特定の世帯だけ抽出

2000年		2005年					両親と 独身の子	世帯主 親
%	世帯数	ShortPat	世帯数	増減数	%	順位		
11.3	54,578	父母伴	54,707	129	10.8	4		
9.8	47,131	父母好	46,952	-179	9.2	5		
9.4	45,122	父母伴好	42,080	-3,042	8.3	6		
合計	146,831		143,739	-3,092				

2000年		2005年					母子世帯	
%	世帯数	ShortPat	世帯数	増減数	%	順位	二人世帯 とは限らない	世帯主 母
2.0	9,746	母伴	11,474	1,728	2.3	7		
1.9	8,921	母好	10,118	1,197	2.0	8		
1.0	4,895	母伴好	5,654	759	1.1	11		
合計	23,562		27,246	3,684				

0.8	3,711	男袋	4,075	364	0.8	13	二人世帯	母子世帯 世帯主 子
0.2	1,100	女袋	1,209	109	0.2	25		
0.1	304	男袋弟	313	9	0.1	56	子:複数	
0.0	241	男袋妹	279	38	0.1	61		
0.0	100	男袋姉	98	-2	0.0	101		
0.0	55	女袋妹	69	14	0.0	116		
0.0	29	女袋弟	37	8	0.0	159		
0.0	24	男袋兄	32	8	0.0	174		
0.0	8	女袋姉	17	9	0.0	234		
0.0	9	女袋兄	6	-3	0.0	380		
合計	5,581		6,135	554				

6. 家族構成変数を使って SAS で特定の世帯形態の世帯を抽出する方法

家族構成変数は、各世帯の全員の続柄を漢字一文字で表現して、一つの変数の中に、世帯主を左端に配置し、残りの世帯員を続いて隙間なく配列して作成される。従って、「人間」の目でこの変数を見ると、直ちに各世帯の家族構成が一目で認識できる。先に示した表2-1や表2-2の様な度数分布表の作成は、この家族構成変数なしでは、かなり複雑な計算処理が必要と思われる。一方、家族構成変数を使って処理をする際には、KanjiPat や ShortPat という家族構成変数の中に、2バイトモードの文字(漢字)が詰め込まれているので、その中に含まれているそれぞれの漢字をプログラミングの際、どう認識するのかが問題となる。本システムは SAS データベース上に構築されているので、本節では、SAS 言語を使って、特定の世帯構造の世帯を検索する方法を解説する。

6.1 一組の指定した続柄だけから成る世帯だけを抽出する方法

「指定した続柄だけから成る世帯」を抽出する SAS コードを、図5に示す。同図の A や B は、同一セッションでその前に SAS で実行済みなら不要で、C の部分だけを実行すればよい。或いは、B のマクロ定義だけ取り出して別の独立した SAS プログラムとしてブラックボックス化して保存しておけば、実行時には、B の箇所は **%include** 文1行で置き換えることができる。C で、抽出したい世帯構造に含まれる全ての世帯員の続柄漢字をコンマで区

切る。この例では、「父・母・倅・好だけから成る世帯」を全て抽出している。

***下の①～④の「=」の右に、自分の環境を入力して下さい;**

***①: %let DS_folder_name = **D:¥国調匿名¥匿名sasDS**; *←★検索対象のSASデータセットが保存されているfolder名; A

***②: %let sasDS = **HHpat2000**; *←★検索対象のSASデータセット名;

***③: %let result_dataset=**resultDS1**; *←★検索結果が保存されるテンポラリSASデータセット名(★何でも構わない);

***④: %let title=**国調・匿名2000年**; *←★結果ビューア画面に明記する; *★なくてもよい;

%macro get_only_household (t1,t2,t3,t4,t5,t6,t7,t8); *ブラックボックス化;

libname kokutoku "&DS_folder_name";

data temp ; set kokutoku.&sasDS; xshortPat=shortPat; run; B

data temp1; set; find=INDEX(xshortPat,"&t1"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp1"; run;

data temp2; set; find=INDEX(xshortPat,"&t2"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp2"; run;

data temp3; set; find=INDEX(xshortPat,"&t3"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp3"; run;

data temp4; set; find=INDEX(xshortPat,"&t4"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp4"; run;

data temp5; set; find=INDEX(xshortPat,"&t5"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp5"; run;

data temp6; set; find=INDEX(xshortPat,"&t6"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp6"; run;

data temp7; set; find=INDEX(xshortPat,"&t7"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp7"; run;

data temp8; set; find=INDEX(xshortPat,"&t8"); if find then do; substr(xshortPat,(find-1)*2+1,2)=" "; output; end; run; proc print; title "DS=temp8"; run;

data &result_dataset(drop=xshortPat find); set temp8; run; *上の③で指定したSASデータセットに検索結果が保存される;

%mend get_only_household;

****★マクロ変数 t1～t8を指定して下さい。**

***★必要な続柄(最大8個)を列挙する。★列挙した全ての続柄★だけ★から成る世帯を抽出する★;** C

*****⑤: %get_only_household (母,倅,父,好, , , ,); *マクロを実行する。続柄漢字の順序は問わない。**

図5 指定した続柄だけから成る世帯を抽出する SAS コード

表3 2000年国調・匿名データ：「父母倅好」だけから成る世帯の構成と戸数

順位	kanjiPat	頻度	%	累積%	順位	kanjiPat	頻度	%	累積%
1	父母倅好	16,114	35.71	35.71	38	父母好倅倅好好	5	0.01	99.8
2	父母好倅	15,426	34.19	69.9	39	父母好倅倅好倅	5	0.01	99.81
3	父母倅倅好	2,367	5.25	75.15	40	父母好倅倅倅倅	5	0.01	99.82
4	父母好好倅	2,029	4.5	79.64	41	父母倅好好好倅	5	0.01	99.83
5	父母倅好好	1,921	4.26	83.9	42	父母倅好好倅倅	5	0.01	99.84
6	父母倅好倅	1,918	4.25	88.15	43	父母倅倅好倅好	5	0.01	99.86
7	父母好倅倅	1,841	4.08	92.23	44	父母倅倅倅好倅	5	0.01	99.87
8	父母好倅好	1,760	3.9	96.13	45	母父倅好倅	5	0.01	99.88
9	父母倅倅倅好	140	0.31	96.44	46	母父倅倅好	5	0.01	99.89
10	父母好好倅倅	131	0.29	96.73	47	父母好好好倅好			
11	父母好倅倅好	118	0.26	96.99	48	父母倅好好好好			
12	父母好好好倅	114	0.25	97.25	49	父母倅好好倅好			
13	父母倅倅好好	106	0.23	97.48	50	父母倅好倅倅倅			
14	父母好倅倅倅	104	0.23	97.71	51	父母好好倅好好			
15	父母倅好倅倅	101	0.22	97.93	52	父母倅倅好好倅			
16	父母好倅好好	100	0.22	98.16	53	母父好倅好			
17	父母倅好好倅	100	0.22	98.38	54	父母好好倅好倅			
18	父母好倅好倅	99	0.22	98.6	55	父母好倅好好倅			
19	父母倅好好好	97	0.21	98.81	56	父母好倅好倅好			
20	父母好好倅好	91	0.2	99.01	57	父母倅倅好好好			
21	父母倅倅好倅	91	0.2	99.22	58	父母好好好倅好倅			
22	父母倅好倅好	79	0.18	99.39	59	父母好好好倅倅			
23	母父倅好	48	0.11	99.5	60	父母好好倅好倅倅			
24	母父好倅	41	0.09	99.59	61	父母好倅好好好			
25	父母倅倅倅好好	11	0.02	99.61	62	父母好倅倅好倅倅			
26	母父好倅倅	9	0.02	99.63	63	父母倅好好好倅倅			
27	父母好好好好倅	8	0.02	99.65	64	父母倅好倅好倅			
28	母父倅好好	8	0.02	99.67	65	父母倅好倅倅好			
29	父母好倅好倅倅	7	0.02	99.68	66	父母倅倅好倅好好			
30	父母倅倅好倅倅	7	0.02	99.7	67	父母倅倅倅好倅			
31	母父好好好	7	0.02	99.71	68	母父好好好倅			
32	父母好倅倅倅好	6	0.01	99.73	69	母父好好好好			
33	父母倅好倅好好	6	0.01	99.74	70	母父好好倅倅			
34	父母倅好倅倅好	6	0.01	99.75	71	母父好倅好好			
35	父母倅倅倅倅好	6	0.01	99.77	72	母父倅好好倅			
36	父母好好好倅倅	5	0.01	99.78	73	母父倅倅好倅			
37	父母好好好倅倅	5	0.01	99.79	74	母父倅倅倅好			
	合計	45,122	100.00						100.00

頻度は5未満は非表示

図5のプログラムのBの部分は、一見、data stepで同じことを8回(国調・匿名データでは、1世帯当たりの世帯員数を8人までに限定)繰り返しているように見えるが、各data stepを実行する毎に対象世帯が絞られて行くので、後続のdata stepでは、その都度、その直前に作成されたデータセットに対して、新たな検索をかけている。つまり、data stepが一つずつ実行されるにつれて、新たに絞り込まれたデータセットに対して、検索処理が行われる仕組みになっており、最終的には、指定された続柄だけの世帯が抽出される。「set 文」が、処理対象のデータセット名の明示なしで、「set;」だけの時は、直前で作成、又は使用したデータセットが自動的に処理の対象になるSASのシステム仕様を利用している。

Bの処理では、SAS関数のKINDEXとSUBSTRを使って、家族構成変数に含まれる2バイトモードの漢字から成る文字変数にアクセスして識別処理をしている。

実行結果のSASデータセットから世帯構造のパターンの頻度分布を作成し、出現頻度順に並べ替えた表を表3に示す³。

6.2 複数の異なる世帯構造の世帯をまとめて抽出する方法

複数の異なる世帯構造の世帯をひとまとめにして分析したいこともある。その場合には、6.1節に示したアルゴリズムを、複数回実行してから、それらを縦にマージすればできるが、本節では、その処理をひとつのSASプログラムでできるようにとりまとめて、複数のデータセットを作成した後、マージ・ソートをすることで実現している(図6)。

図6に示す基本的なアルゴリズムは図5で採用している方法と基本的には同じである。この例では、「世帯主の両親と独身の子」の世帯を抽出するために、家族構成変数ShortPatが「父母俣」、「父母好」、「父母俣好」である3つの異なる世帯構造の世帯を順次特定した後、一つのデータセットにマージしている。そのデータセットを作成した後、世帯id番号順にソートしている。

7. 夫婦と子供から成る世帯

国調の調査票情報や匿名データには、「世帯の家族類型」という変数が組み込まれており、この変数には、「2:夫婦と子供から成る世帯」という世帯形態があることは第4節でも述べた。2000年と2005年の匿名データからこの家族類型と家族構成変数(短縮版ShortPat)のクロス表を作成した(表4)。この表を見ると、普通一般に「夫婦と子供から成る世帯」からイメージする世帯形態とはかなり印象が異なることが分かる。ただ、2000年と2005年の2回の調査とも、この「家族類型」の中では、「両親と独身の子」から成る「父母好」、「父母俣」、「父母俣好」が世帯形態の上位3位で大多数を占めることが確認できた。それに続く戸数の多い世帯形態では、「父母俣」や「父母娘」の様に、「世帯主の両親と同居する婚姻中の子」の世帯があり、嫁や婿が別の世帯にいることを表している。

更に、「男爺婆」(独身の息子が世帯主で、その両親と同居)もかなりの戸数があり、この世

³ 家族構成変数のKanjiPatでは、一番左端の世帯員が世帯主なので、両親のどちらが世帯主か判別できる。因みに、短縮版の家族構成変数ShortPatでは、便宜上、夫婦のどちらが世帯主でも、「父母」の順に表記を統一している。

帯に更に同居する世帯主の弟・妹がいる世帯、「男爺婆弟」と「男爺婆妹」を加算すると、ある程度、纏まった戸数になっている。さすがに、「男爺婆姉」(両親と同居する独身男性の世帯主の姉も同居)の世帯数は5戸未満しかない。

表4 ShortPat と「世帯の家族類型」(2.夫婦と子供から成る世帯)のクロス表

国調・匿名データ2000年				国調・匿名データ2005年				
shortPat	familyType			shortPat	familyType			
	2:夫婦と子供から成る世帯 度数				2:夫婦と子供から成る世帯 度数			
1 奥爺袋			28 父母娘	223	1 奥爺袋	6	28 亭爺袋	32
2 奥舅姑	6		29 父母娘好	35	2 奥舅姑		29 亭爺袋姉	
3 女爺袋	74		30 父母倅	314	3 女爺袋弟		30 亭爺袋弟	
4 女爺袋兄			31 父母倅好	41	4 女袋爺妹		31 亭爺袋妹	
5 女爺袋弟			32 父母倅娘	11	5 女爺袋	81	32 亭舅姑	
6 女爺袋妹			33 父母倅倅	32	6 女爺袋姉		33 父母嫁	52
7 女舅姑			34 父母倅倅	10	7 女爺袋戚		34 父母好	46,952
8 男袋爺姉弟			35 父母倅倅		8 女爺袋弟		35 父母好戚	39
9 男袋爺妹弟			36 父母倅	54,578	9 女袋爺妹	5	36 父母婿	6
10 男爺袋	509		37 父母倅嫁		10 女舅姑		37 父母娘	265
11 男爺袋兄			38 父母倅好	45,122	11 男袋爺姉		38 父母娘好	27
12 男爺袋姉			39 父母倅好	24	12 男袋爺弟		39 父母娘戚	
13 男爺袋戚			40 父母倅戚	49	13 男袋爺妹		40 父母倅	433
14 男爺袋弟	33		41 父母倅娘	28	14 男爺袋	468	41 父母倅好	39
15 男爺袋弟妹			42 父母倅娘	6	15 男爺袋兄	5	42 父母倅戚	
16 男爺袋妹	44		合計	148,452	16 男爺袋姉	6	43 父母倅婿	
17 男爺袋妹弟					17 男爺袋姉戚		44 父母倅娘	
18 男舅姑					18 男爺袋姉弟		45 父母倅倅	45
19 亭爺袋	35				19 男爺袋姉妹		46 父母倅倅	
20 亭爺袋弟					20 男爺袋戚		47 父母倅	54,707
21 亭舅姑					21 男爺袋弟	44	48 父母倅嫁	
22 父母嫁	36				22 男爺袋弟戚		49 父母倅好	42,080
23 父母好	47,131				23 男爺袋弟妹		50 父母倅好	20
24 父母好嫁					24 男爺袋妹	23	51 父母倅戚	56
25 父母好戚	47				25 男爺袋妹弟		52 父母倅婿	
26 父母好婿					26 男舅姑	12	53 父母倅娘	40
27 父母婿	15				27 男舅姑弟		54 父母倅娘	6
							合計	145,499

参考までに、2000年の国調・匿名データに対して、図6(次頁)のSASプログラムを実行して出力されたデータセットを使って、世帯の全員が見える全体版家族構成変数(KanjiPat)で頻度表を作成し、頻度順にソートした。(表5)⁴。両親と独身の子が一人の世帯(1位と2位の合計)が66,221戸で、全体(146,831戸)の約45%を占めている。次いで、両親と独身の子が二人の世帯(3~6位)を合わせた61,878戸で続く。これら上位6位までの合計が128,099戸で、全体の87%を占めている。また、2000年国調・匿名データには、全戸数482,004戸あるので、「両親と独身の子が二人の世帯」は12.8%にしか過ぎない。我が国では、「夫婦と子供2人の世帯」が標準世帯として、これまで長く国の統計や税金の試算などにおいて使われているが、少子高齢化が進行している現代では、標準でもモデルでもなくなってしまっている。「子供」の年齢を20歳未満に限定すれば、更に戸数は少なくなる。

次に、同様に、2005年データでも実行して、2000年と2005年の「両親と独身の子供二

⁴ 図5のプログラムを実行して作成された表3は、「世帯主の両親が独身の息子と娘だけと同居している」世帯、図6のプログラムの実行結果の表5は「世帯主の両親が独身の息子 AND/OR 独身の娘と同居している」世帯である。両表とも全体版の家族構成変数 KanjiPat を使っているため、世帯員全員の構成が見える。KanjiPat にある「倅」と「好」の並び順は、提供された元ファイルの順のまま、恐らく子供の年齢順に並んでいると思われる。

人の世帯」の世帯数を比較した(表6:次頁)。父親が世帯主の世帯は減少傾向、逆に母親が世帯主の世帯が増加傾向にあり、かつ、その両方の合計も減少していることが読み取れる。

図6 複数の異なる世帯構造の世帯をまとめて抽出する SAS コード

```

*☆国調・匿名データ専用プログラム(世帯あたり世帯員数8人まで)☆;
*☆☆下の①~②の「=」の右に、自分の環境を入力して下さい☆☆;
*★★ ; %let year=2000; *←★調査年次を西暦で記入する;
*★★①; %let DS_folder_name = D:\国調匿名\匿名sasDS; *←★検索対象のSASデータセットの保存folder名;
*★★②; %let sasDS = xHHpat&year; *←★検索対象のSASデータセット名;
*★★③; %let result_dataset= resultDS1; *←★検索結果が保存されるテンポラリSASデータセット名(★何でも構わない);
*★★④; %let title=国調・匿名&year.年; *←★結果ビューア画面に明記する;

%macro get_only_household_many (No,t1,t2,t3,t4,t5,t6,t7,t8);
*☆macro変数:No=検索したい世帯構成の種類に、順に1~8の番号を付ける☆;
*☆macro変数:t1~t8=検索したい家族構成変数を一文字ずつ列挙する☆;
libname kokutoku "&DS_folder_name";
data temp; set kokutoku.&sasDS; xshortpat=shortpat; run;

data temp1; set; find=KINDEX(xshortpat,"&t1"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp2; set; find=KINDEX(xshortpat,"&t2"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp3; set; find=KINDEX(xshortpat,"&t3"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp4; set; find=KINDEX(xshortpat,"&t4"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp5; set; find=KINDEX(xshortpat,"&t5"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp6; set; find=KINDEX(xshortpat,"&t6"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp7; set; find=KINDEX(xshortpat,"&t7"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;
data temp8; set; find=KINDEX(xshortpat,"&t8"); if find then do; substr(xshortpat,(find-1)*2+1,2)=" "; output; end;

data HHstructure&No.; set temp8; run; *★冒頭の③で指定したテンポラリSASデータセットに検索結果が保存される;
proc freq; tables shortPat / norow nocol nopercnt; title "途中結果:data=HHstructure&No.:shortPat"; run;
proc freq; tables kanjiPat / norow nocol nopercnt; title "途中結果:data=HHstructure&No.:kanjiPat"; run;
%mend;

*★空のdatasetを事前に作成しておく;
data HHstructure1 HHstructure2 HHstructure3 HHstructure4 HHstructure5 HHstructure6 HHstructure7 HHstructure8; run;
*☆☆以下のマクロ変数を指定して下さい。この例では、父母と同居する独身の「息子が娘、又は、その両者」だけが同居する世帯を検索するので、以下の⑤~⑦だけを指定し、⑧以下はそのまま放置する;
*1★★⑤; %get_only_household_many(1,父,母,伴,,,,); *←必要な続柄(最大8個)を列挙する ;
*2★★⑥; %get_only_household_many(2,父,母,好,,,,); *←必要な続柄(最大8個)を列挙する ;
*3★★⑦; %get_only_household_many(3,父,母,伴,好,,,,); *←必要な続柄(最大8個)を列挙する ;
*4★★⑧; %get_only_household_many(4,,,,,,); *←必要な続柄(最大8個)を列挙する ;
*5★★⑨; %get_only_household_many(5,,,,,,); *←必要な続柄(最大8個)を列挙する ;
*6★★⑩; %get_only_household_many(6,,,,,,); *←必要な続柄(最大8個)を列挙する ;
*7★★⑪; %get_only_household_many(7,,,,,,); *←必要な続柄(最大8個)を列挙する ;
*8★★⑫; %get_only_household_many(8,,,,,,); *←必要な続柄(最大8個)を列挙する ;

*☆☆上のSASマクロを実行すると、複数個の検索結果のデータセット(HHstructure1,HHstructure2,...)が作成されるので、次のdata stepで、冒頭の%let文で指定したtemporary SASデータセットにまとめて保存される☆;
data &result_dataset;
set HHstructure1 HHstructure2 HHstructure3 HHstructure4 HHstructure5 HHstructure6 HHstructure7 HHstructure8;
proc sort; by HHidNo; run;
*★★★★★★この段階で、検索したい世帯の抽出は完了した。★★★★★★;

```

このマクロ定義の箇所は、変更禁止

このマクロ実行箇所は、⑤~⑫を適宜、33個の漢字を使って変更する。

8. 母親と子から成る世帯

多くの母子世帯は、家計消費の面で苦しい生活を強いられている。その実態を分析するには、全国消費実態調査⁵の収支情報を使う必要があるが、その点は現在作業中であり、本稿では、対象となる世帯の分布を調べるに留める。

まず、母親の現在の婚姻状況を考慮せずに、「母親と独身の子から成る世帯」の戸数を、2000年と2005年の国調・匿名データから数えた(表7：次頁)。その5年間で、3,684世帯増加している。2000年時点の総世帯数が23,562⁶なので、15.6%の増加である。国調はその後、2010年、2015年、2020年と実査されているので、仮に、調査時点毎に15%ずつ増加していくとすれば、2020年の時点では、およそ4万世帯ほどに膨れあがっていると推定される。

次に、「母子世帯」について分析する。普通一般的に「母子世帯」と言えば、世帯主は母親で、現在婚姻中ではない世帯を連想する(表8：次頁)。2000年の時点で19,492世帯、2005年で22,790世帯なので、その5年間で約17%の増加になっており、この増加率が続けば、直近の国勢調査(2020年実査)では、約34,000世帯ほどに増加していることが予想される。

表7、表8をよく見ると、男子の「子」がいる世帯の方が、女子がいる世帯よりも世帯数が多いことに気づく。結婚しない男子の方が多いのか、あるいは男子の方が、婚期が遅いということなのかどうかは、今後の分析で解明したい。

表5 2000年国調・匿名データ「両親と独身の子から成る世帯」：頻度順

順位	kanjiPat	世帯数	%	累積%
1	父母倅	35,765	24.4	24.4
2	父母好	30,456	20.7	45.1
3	父母倅好	16,114	11.0	56.1
4	父母倅倅	15,980	10.9	67.0
5	父母好倅	15,426	10.5	77.5
6	父母好好	14,358	9.8	87.2
7	父母倅倅倅	2,436	1.7	88.9
8	父母倅倅好	2,367	1.6	90.5
9	父母好好倅	2,029	1.4	91.9
10	父母好好好	2,004	1.4	93.3
11	父母倅好好	1,921	1.3	94.6
12	父母倅好倅	1,918	1.3	95.9
13	父母好倅倅	1,841	1.3	97.1
14	父母好倅好	1,760	1.2	98.3
15	母父倅	163	0.1	98.4
16	父母倅倅倅倅	154	0.1	98.5
17	母父好	144	0.1	98.6
18	父母倅倅倅好	140	0.1	98.7
19	父母好好倅倅	131	0.1	98.8
20	父母好好好好	118	0.1	98.9
21	父母好倅倅好	118	0.1	99.0
22	父母好好好倅	114	0.1	99.1
23	父母倅倅好好	106	0.1	99.1
24	父母好倅倅倅	104	0.1	99.2
25	父母倅好倅倅	101	0.1	99.3
26	父母好倅好好	100	0.1	99.3
27	父母倅好好倅	100	0.1	99.4
世帯数100未満の28位～92位(最下位)略				
合計		146,831	100.0	

表6 2000年から2005年の「両親と独身の子供2人の世帯」の戸数の変遷

国調・匿名2000年：標準世帯			国調・匿名2005年：標準世帯					
	kanjiPat	世帯数	%		kanjiPat	世帯数	%	
1	父母倅好	16,114	26.0	↓	1	父母倅倅	15,472	26.1
2	父母倅倅	15,980	25.7		2	父母倅好	15,225	25.7
3	父母好倅	15,426	24.9		3	父母好倅	14,543	24.6
4	父母好好	14,358	23.1		4	父母好好	13,726	23.2
5	母父倅倅	60	0.1	↑	5	母父倅倅	69	0.1
6	母父倅好	48	0.1		6	母父好倅	67	0.1
7	母父好倅	41	0.1		7	母父倅好	60	0.1
8	母父好好	38	0.1		8	母父好好	40	0.1
計		62,065	100.0		計		59,202	100.0

⁵ 令和元年から「全国家計構造調査」と調査名を変更して実施されている。

⁶ 国調・匿名データには復元乗率の変数がないので、本稿ではそのままの件数を使っている。国調・匿名データは1%抽出しているので、実際の件数はこれらの値を100倍すればおよその数字が推定できると思う。

表7 母親(婚姻状況問わず)と独身の子から成る世帯

2000年		2005年			子の人数
kanjiPat	世帯数	kanjiPat	世帯数	増減数	
1 母 伴	7,557	母 伴	8,975	1,418	1
2 母 好	6,952	母 好	7,948	996	
3 母 伴 伴	1,922	母 伴 伴	2,216	294	2
4 母 伴 好	1,815	母 伴 好	2,105	290	
5 母 好 好	1,723	母 好 好	1,919	196	
6 母 好 伴	1,630	母 好 伴	1,876	246	
7 母 伴 伴 伴	244	母 伴 伴 伴	268	24	3
8 母 好 好 好	230	母 好 好 伴	267	54	
9 母 伴 伴 好	223	母 伴 伴 好	260	37	
10 母 好 伴 伴	220	母 伴 好 好	246	32	
11 母 伴 好 好	214	母 好 好 好	237	7	
12 母 好 好 伴	213	母 好 伴 伴	237	17	
13 母 伴 好 伴	210	母 伴 好 伴	220	10	
14 母 好 伴 好	181	母 好 伴 好	196	15	
1~14位：合計	23,334	1~14位：合計	26,970	3,636	
15~59位：合計 (世帯員5名以上)	228	15~59位：合計 (世帯員5名以上)	276	48	
総数	23,562	総数	27,246		

表注：本表には現在婚姻中の母親も含まれており、その配偶者は単身赴任や病院・施設等の別世帯に生活していると思われる。

表8 母子世帯:「独身の母(世帯主)」と独身の子から成る世帯

2000年		2005年			子の人数
kanjiPat	世帯数	kanjiPat	世帯数	増減数	
1 母 伴	6,681	母 伴	7,936	1,255	1
2 母 好	6,028	母 好	6,928	900	
3 母 伴 伴	1,519	母 伴 伴	1,768	249	2
4 母 伴 好	1,358	母 伴 好	1,597	239	
5 母 好 好	1,278	母 好 伴	1,464	246	
6 母 好 伴	1,218	母 好 好	1,458	180	
7 母 伴 伴 伴	173	母 伴 伴 伴	208	35	3
8 母 伴 伴 好	161	母 好 好 伴	198	56	
9 母 好 伴 伴	160	母 伴 伴 好	185	24	
10 母 伴 好 伴	159	母 伴 好 好	178	32	
11 母 好 好 好	156	母 好 伴 伴	177	17	
12 母 伴 好 好	146	母 好 好 好	169	13	
13 母 好 好 伴	142	母 伴 好 伴	150	-9	
14 母 好 伴 好	126	母 好 伴 好	142	16	
1~14位：合計	19,305	1~14位：合計	22,558	3,253	
15~59位：合計 (世帯員5名以上)	187	15~59位：合計 (世帯員5名以上)	232	45	
総数	19,492	総数	22,790		

次に、表7に示す母子世帯の中で、「母親(婚姻状況問わず)と独身の子一人」から成る母子世帯(二人世帯:家族構成変数 KanjiPat が「母伴」と「母好」)だけを対象に、調査年別に、2000年(14,509世帯)と2005年(16,923世帯)の「母子の年齢⁷の組み合わせ」を表9に示す。比較的頻度が多いセルを着色している。

2000年と2005年の出現頻度を比較すると、2000年では、母親が42歳~62歳で、独身の子が17歳~37歳の組み合わせに集中しているが、2005年では、母親が47歳~77歳で、独身の子が17歳~52歳の組み合わせに広がっていることが観察できる。要するに、母子ともに高齢化の傾向にあると言える。

表9 母親(婚姻状況問わず)と独身の子一人から成る母子世帯

母親と独身の子から成る母子世帯(二人世帯):母親年齢X独身の子年齢(国調・匿名データ2000年)

母親年齢	独身の子:年齢															合計	
	2	7	12	17	22	27	32	37	42	47	52	57	62	67	72		77
17	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
22	119	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	130
27	245	208	9	0	0	0	0	0	0	0	0	0	0	0	0	0	462
32	181	■	132	■	■	■	■	■	■	■	■	■	■	■	■	■	623
37	97	196	277	123	■	■	■	■	■	■	■	■	■	■	■	■	697
42	24	91	231	406	80	■	■	■	■	■	■	■	■	■	■	■	835
47	1	27	193	574	602	169	■	■	■	■	■	■	■	■	■	■	1,569
52	0	7	96	324	800	919	104	■	■	■	■	■	■	■	■	■	2,252
57	0	0	11	93	275	838	594	85	■	■	■	■	■	■	■	■	1,897
62	0	0	0	18	85	357	606	473	95	6	■	■	■	■	■	■	1,640
67	0	0	0	0	12	105	267	516	332	94	8	■	■	■	■	■	1,334
72	0	0	0	0	0	12	79	202	341	313	125	5	■	■	■	■	1,077
77	0	0	0	0	0	0	5	38	165	311	328	54	6	■	■	■	907
82	0	0	0	0	0	0	0	0	35	148	235	169	45	■	■	■	636
87	0	0	0	0	0	0	0	0	0	36	119	125	82	59	16	■	442
合計	675	846	949	1,542	1,858	2,403	1,658	1,318	973	908	815	353	133	61	16	■	14,509

凡例:年齢「17」は、「15歳~19歳」を示す。「独身」は、未婚、死別、離別を含む。

母親と独身の子から成る母子世帯(二人世帯):母親年齢X独身の子年齢(国調・匿名データ2005年)

母親年齢	独身の子:年齢															合計	
	2	7	12	17	22	27	32	37	42	47	52	57	62	67	72		77
17	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
22	127	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	139
27	189	208	9	0	0	0	0	0	0	0	0	0	0	0	0	0	406
32	209	408	167	13	0	0	0	0	0	0	0	0	0	0	0	0	797
37	141	334	358	143	■	■	■	■	■	■	■	■	■	■	■	■	980
42	45	140	343	393	128	■	■	■	■	■	■	■	■	■	■	■	1,050
47	3	41	191	578	455	109	■	■	■	■	■	■	■	■	■	■	1,377
52	0	7	86	309	679	645	159	5	■	■	■	■	■	■	■	■	1,890
57	0	0	9	131	373	959	790	150	■	■	■	■	■	■	■	■	2,413
62	0	0	0	7	82	282	741	644	97	■	■	■	■	■	■	■	1,857
67	0	0	0	0	14	73	309	598	520	100	5	■	■	■	■	■	1,619
72	0	0	0	0	0	11	95	286	556	470	126	12	■	■	■	■	1,556
77	0	0	0	0	0	0	10	77	185	404	402	131	■	■	■	■	1,211
82	0	0	0	0	0	0	0	5	35	157	347	336	50	8	0	■	938
87	0	0	0	0	0	0	0	0	0	26	122	242	193	71	27	■	684
合計	720	1,150	1,163	1,574	1,735	2,080	2,104	1,765	1,395	1,161	1,002	721	245	79	27	■	16,923

10. 今後の展望

本論文では、家族構成変数を使って、分析したい世帯構造の世帯を抽出するプログラミング上の技法を解説した後、その計算結果の解釈を行った。ここで示した全ての分析表は、第2節の図2のデータセットがあれば、分析したい世帯構造の世帯を直ぐに検索することができる。

⁷ 国調・匿名データでは、世帯員の年齢は、5歳刻みの階級値で表現されている。各世帯員の年齢を実数値で表現するために、各年齢階層の中間値を採用している。例えば、年齢階級値が1は、「0歳~4歳」を意味するので、この階級の属する世帯員の年齢実数値を(0+4)÷2=2歳とした。なお、最高齢の年齢階級18は「85歳以上」を意味するが、便宜上、「85歳~90歳」として、87歳とした。尚、表9で、■のセルは頻度が1~4を意味する。

ここ数年に亘って家族構成変数を開発してきたが、実際に使ってみた過程で、細かいバージョンアップをしてきた。設計段階では、「世帯主の子」を分類するのに必要な漢字表現は、性別と配偶者の有無さえ区別できれば十分と考えていた。しかし、第8節で議論した母子世帯を考えた時、現在「独身」の子は、未婚なのか、結婚歴有りかを区別しないと、綿密な議論ができないことに気がついた。もちろん、今のままでも、図2に示した「世帯員続柄データセット」の中には、各世帯員の詳しい婚姻状況を表す変数が含まれているので、その区別は可能であり、事実、第8節の分析表はその変数を加味することで出力できた。ただ、分析をする度に煩わしい余分な計算処理が必要となり、これは当初意図したことではなかった。本文中でも触れたが、現在実装している2種類の家族構成変数の他に、更に、上で述べた「子の現在の婚姻状況」を、「1:未婚」、「2:婚姻中(事実婚含む)」、「3:結婚歴あるが現在独身」の3種類に分類できる第三の家族構成変数を新設する予定⁸である。世帯主の子以外の続柄の家族については、婚姻状況が問題になった際に、その都度、図2の「世帯員続柄データセット」に含まれている婚姻状況を示す変数 marry1~marry8 を参照すればいいと判断している。

現在我々が関心を持っているテ

表10 匿名データの調査年次

ーマのひとつに、異なる統計調査のデータを「世帯構成」毎に結合させることによって、調査環境が厳しくなっている現代

匿名データ	調査年次					
	国勢調査			2000	2005	2010
全国消費実態調査	1989	1994	1999	2004	2009	2014
就業構造基本調査	1992	1997	2002	2007	2012	2017

社会において、入手できたデータを最大限に活用したいという意図がある。小規模ながら、国調と全国消費実態調査の匿名データを使って、この二つの調査結果を結合する実験結果を昨年の本研究会で一部を披露したが、まだまだ開発途上にある。5年周期で実施されるこの二つの統計調査の実施年は、表10に示すように、1年ずれているだけで、結合してもさほど問題があるとは考えていない。

更に、就業構造基本調査も5年周期で、国調の2年後に実施されているので、こちらの調査データともリンクすれば、それぞれ単独で分析するよりは、多くの情報が入手できると確信している。いずれの場合も、家族構成毎に、3つの異なる統計調査のデータを統合して分析を進める予定である。当分は匿名データで実験を継続して、その後、調査票情報に切り替えた実験を行い、分析手法を確立した後、将来的には、オンサイト利用で同様の分析を行って、より正確な分析結果を得ることを目指している。

⁸ 世帯主の場合は、現行のままでも、図4(第3節)に示すように、「父・母」、「亭・奥」、「夫・婦」、「男・女」として区別している。ただし、最後の「男・女」の表現は、「現在独身」を意味しており、これだけでは結婚歴は分からない。また、「父・母」の場合も、父母が揃っているなら、現在婚姻中だと分かるが、表7や表8で示したように、「父」又は「母」だけなら、婚姻状況はそれだけでは決めつけられない。いずれの場合も、変数 marry1(世帯主の婚姻状況の変数)を参照すれば簡単に分かることなので、このままで不便はないと考えている。

参考文献

- ①伊藤彰彦・中川雅義・周防節雄・米澤香・安井浩子・新井郁子（2017）『新世帯類型の構築と世帯構造の変動に係る長期時系列分析の結果』、(公財)統計情報研究開発センター、2017年
- ②周防節雄・安井浩子(2020a)「国勢調査の続柄情報に婚姻状況・性別・年齢を加味して世帯員構成を直感的に表現できる変数の開発」、『官民オープン利活用の動向及び人材育成の取り組み(2019年度)報告要旨集』、独立行政法人統計センター、pp.81-91、2020年3月
- ③周防節雄(2020b)「世帯を調査客体とする公的マイクロデータの世帯構造を直感的に把握できる「家族構成変数」の開発と利用(1)」、『ESTRELA 9月号』、pp.25-30、(公財)統計情報研究開発センター、2020年9月
- ④周防節雄(2020c)「世帯を調査客体とする公的マイクロデータの世帯構造を直感的に把握できる「家族構成変数」の開発と利用(2)」、『ESTRELA 11月号』、pp.25-31、(公財)統計情報研究開発センター、2020年11月
- ⑤周防節雄(2021)「国勢調査のマイクロデータから新開発した家族構成変数と、それを利用した世帯構造の分析事例の紹介」、『官民オープン利活用の動向及び人材育成の取り組み(2020年度)報告要旨集』、独立行政法人統計センター、pp155-167、2021年3月
- ⑥周防節雄(2022a)「国勢調査と全国消費実態調査の匿名データの統計量を家族構成変数で統合する実験結果の紹介」、独立行政法人統計センター、2022年3月
https://www.nstac.go.jp/sys/files/211118_8.pdf
- ⑦周防節雄(2022b)「国勢調査と全国消費実態調査の匿名データの統計量を家族構成変数で統合する試み」、『ESTRELA 4月号』、pp.2-9、(公財)統計情報研究開発センター、2022年4月

謝辞

本論文の作成に際しては、国勢調査の匿名データを使用したが、その際、(株)SAS インステチュート ジャパンから使用料を援助して頂いた。また、匿名データの利用に際しては、独立行政法人統計センターから便宜を図って頂いた。ここに記して謝意に替えたい。