

国勢調査マイクロデータから 全自動で構築した SASデータベースに基づく 新世帯類型の設計と開発

安井 浩子、中川雅義

(公財)統計情報研究開発センター
研究開発部

周防 節雄

兵庫県立大学・名誉教授

2016/07/22

SASユーザー総会2016

@神戸国際会議場

国勢調査における世帯単位の 従来の家族類型

- 個人の消費には、
 - ・ 自分自身の消費
 - ・ 家族のための消費
 - ・ 家族と一緒にいるための消費などがある。



個人が属する世帯の構造も消費動機に少なからず影響を及ぼしている。

- 世帯の構造別世帯数として、昭和45年国勢調査より「家族類型別」世帯数を公表

➤ 従来の家族類型

平成22年国勢調査用語の解説より作成

区分	内訳
親族のみの世帯	二人以上の世帯員から成る世帯のうち、世帯主と親族関係にある世帯員のみからなる世帯
核家族世帯	(1) 夫婦のみの世帯
	(2) 夫婦と子供から成る世帯
	(3) 男親と子供から成る世帯
	(4) 女親と子供から成る世帯
その他の親族世帯	(5) 夫婦と両親から成る世帯
	(6) 夫婦とひとり親から成る世帯
	(7) 夫婦、子供と両親から成る世帯
	(8) 夫婦、子供とひとり親から成る世帯
	(9) 夫婦と他の親族(親、子供を含まない)から成る世帯
	(10) 夫婦、子供と他の親族(親を含まない)から成る世帯
	(11) 夫婦、親と他の親族(子供を含まない)から成る世帯
	(12) 夫婦、子供、親と他の親族から成る世帯
	(13) 兄弟姉妹のみから成る世帯
	(14) 他に分類されない世帯
非親族を含む世帯	二人以上の世帯員から成る世帯のうち、世帯主と親族関係にない人がいる世帯
単独世帯	世帯人員が一人の世帯

世帯内の最も若い夫婦と他の世帯員との関係によって世帯を分類

➤ 今回開発した新世帯類型

【一例】 家族類型 「夫婦と子供から成る世帯」

↓ (2分割)

新世帯類型 「世帯主夫婦と子がいる世帯」

「世帯主と両親がいる世帯」

一部抜粋

世帯構造を詳細に分類できる

家族類型	I 核家族世帯			
	夫婦のみ の世帯 (1)	夫婦と子供 から成る 世帯 (2)	男親と子 供から成 る世帯 (3)	女親と子 供から成 る世帯 (4)
A 親族世帯	10,253,150	14,438,390	666,780	3,867,060
(48) 世帯主夫婦と子供がいる世帯	—	14,361,230	—	—
(50) 世帯主と子供がいる世帯	—	—	630,450	3,127,900
(53) 世帯主と両親がいる世帯	—	77,160	—	—
(54) 世帯主とひとり親がいる世帯	—	—	36,330	739,160
(55) 世帯主夫婦がいる世帯	10,253,150	—	—	—

国勢調査の続柄

➤ 利用データ

国勢調査の調査票情報(マイクロデータ)

1980、1990、2000、2010年(大規模調査)

	レコード数
1980年	25,288,436
1990年	17,070,017
2000年	17,571,350
2010年	14,695,119

➤ 続柄

4年分で共通する続柄に再編成した。

次以降で説明

1980年(昭和55年)		
符号		変更後の符号
01	世帯主又は代表者	01
02	世帯主の配偶者	02
03	子	03
04	子の配偶者	04
05	孫	05
06	父母	06
07	祖父母	07
08	兄弟・姉妹	08
09	他の親族	09
10	住み込みの雇人(家事)	10
11	住み込みの雇人(営業)	10
12	その他	11
**	対象外	12

1990年(平成2年)		
符号		変更後の符号
01	世帯主又は代表者	01
02	世帯主の配偶者	02
03	子	03
04	子の配偶者	04
05	孫	05
06	父母	06
07	祖父母	07
08	兄弟姉妹	08
09	他の親族	09
10	家事使用人	10
11	営業使用人	10
12	その他	11
VV	不詳	12

2010年(平成22年)		
符号		変更後の符号
01	世帯主または代表者	01
02	世帯主の配偶者	02
03	子	03
04	子の配偶者	04
05	世帯主の父母	06
06	世帯主の配偶者の父母	06
07	孫	05
08	祖父母	07
09	兄弟姉妹	08
10	他の親族	09
11	住み込みの雇人	10
12	その他	11
VV	不詳(世帯主以外の基本項目 記入不備世帯員)	12
△△	無人調査区	

2000年(平成12年)		
符号		変更後の符号
01	世帯主または代表者	01
02	世帯主の配偶者	02
03	子	03
04	子の配偶者	04
05	世帯主の父母	06
06	世帯主の配偶者の父母	06
07	孫	05
08	祖父母	07
09	兄弟姉妹	08
10	他の親族	09
11	住み込みの雇人	10
12	その他	11
VV	不詳(聞き取り調査世帯)	12

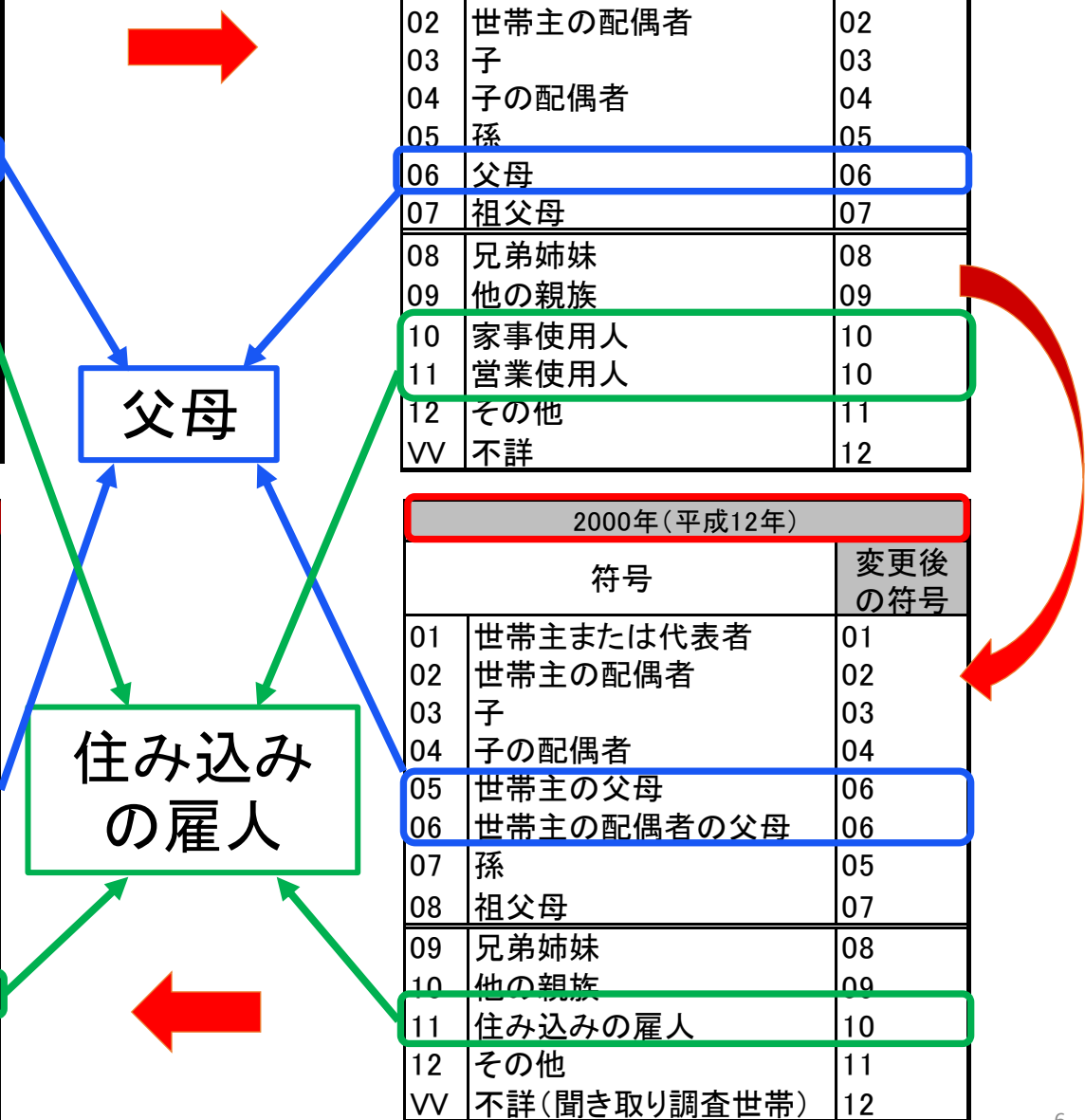


表1 国勢調査の続柄一覧 【注】二重線から上が直系親族



新世帯分類に使用した続柄

01	世帯主又は代表者	直系親族	親族世帯
02	世帯主の配偶者		
03	子		
04	子の配偶者		
05	孫		
06	父母		
07	祖父母		
08	兄弟姉妹	傍系親族	
09	他の親族		
10	住み込みの雇人	非親族	
11	その他		
12	不詳		

国勢調査の調査票情報の全自動 SASデータベース化

➤ 周防(2015)にて、全国消費実態調査の匿名データで実用化済みであり、本システムでも同様の手法を採用

- ① 1回次分だけの「マイクロデータ本体と符号表(メタデータ: Excel形式)を自動的に読み込み、SASデータセットを作成する SASプログラム」を自動生成する SASマクロプログラムを書いた。
このマクロを4回次分実行するだけでデータベースが構築できた。

効率的に行う
ことができた。

例：符号表(1980年の抜粋)

行番号	項目名	階層	位置	バイト数	繰返し	配置	型	小数点	種別	変数名	対象	符号	符号内容	備考
1	調査に関する事項	1												
2	調査名	2	1	2					3	CHOIDT	全	C6	抽出詳細集計(20%)	
3	調査年	2	3	4						YEAR		1980	調査の年(西暦)	
4	調査月	2	7	2						MONTH		10	調査の月	
5	所在地に関する事項	1												
6	市区町村情報	2												
7	地域ブロック符号	3	9	2		3				BLK	全	01	北海道(01北海道)	
8												02	東北(02青森 03岩手 04宮城 05秋田 06山形 07福島)	
...

- ② 新世帯類型コードを設計するために作成されたExcel表をSASプログラムで読み込んで、SASデータベースに組み込んだ後、分析作業を行った。

新世帯類型の作成プロセス

正確にすべての
のパターンを
網羅できる。

- 新世帯類型の設計には、Excel表を活用。

プロセス1: 世帯主を含む直系親族の全ての パターンの析出

- ① 表1(世帯の続柄)にある
直系親族7種類(祖父母、父母、世帯主、世帯主の配偶者、
子、子の配偶者、孫)の全ての組み合わせパターンについて、
有無を「0、1」のパターンで作成。
【注】「父母」だけは「0:いない、1:ひとりだけ、2:両親」の区別を
するため「0、1、2」を使用



7つの続柄の全ての組み合わせパターン
96通り(3×2^5)

- ② 上記「0、1、2」の組み合わせパターンを見ながら、
それぞれの続柄の有無を「続柄の構成」欄に日本語で表現。

有無 パターン #	直系親族							
	70 祖 父母	60 父 母	10 世 帯 主	20 世 帯 主 の 配 偶 者	30 子	40 子 の 配 偶 者	50 孫	
種類	0 1	0 1 2	1	0 1	0 1	0 1	0 1	
1	0	0	1	0	0	0	0	世帯主がいる世帯(=世帯主と傍系親族から成る世帯)
2	0	0	1	0	0	0	1	世帯主と孫がいる世帯
3	0	0	1	0	0	1	0	世帯主と子供の配偶者がいる世帯
4	0	0	1	0	0	1	1	世帯主と子供の配偶者と孫がいる世帯
5	0	0	1	0	1	0	0	世帯主と子供がいる世帯
6	0	0	1	0	1	0	1	世帯主と子供と孫がいる世帯
7	0	0	1	0	1	1	0	世帯主と子供夫婦がいる世帯
8	0	0	1	0	1	1	1	世帯主と子供夫婦と孫がいる世帯
9	0	0	1	1	0	0	0	世帯主夫婦がいる世帯
10	0	0	1	1	0	0	1	世帯主夫婦と孫がいる世帯
11	0	0	1	1	0	1	0	世帯主夫婦と子供の配偶者がいる世帯
12	0	0	1	1	0	1	1	世帯主夫婦と子供の配偶者と孫がいる世帯
13	0	0	1	1	1	0	0	世帯主夫婦と子供がいる世帯
14	0	0	1	1	1	0	1	世帯主夫婦と子供と孫がいる世帯
15	0	0	1	1	1	1	0	世帯主夫婦と子供夫婦がいる世帯
16	0	0	1	1	1	1	1	世帯主夫婦と子供夫婦と孫がいる世帯
17	0	1	1	0	0	0	0	世帯主とひとり親がいる世帯
18	0	1	1	0	0	0	1	世帯主と孫とひとり親がいる世帯
19	0	1	1	0	0	1	0	世帯主と子供の配偶者とひとり親がいる世帯
20	0	1	1	0	0	1	1	世帯主と子供の配偶者と孫とひとり親がいる世帯
	0	1	1	0	1	0	0	世帯主と子供とひとり親がいる世帯
				1	0	1		世帯主と子供と孫とひとり親がいる世帯

プロセス
1②

プロセス
1①

続柄の構成

プロセス
2

新世帯類型用の区分の名称

世帯主がいる世帯(=世帯主と傍系親族から成る世帯)
世帯主と孫がいる世帯
世帯主と子供がいる世帯
世帯主と子供と孫がいる世帯
世帯主と子供がいる世帯
世帯主と子供と孫がいる世帯
世帯主と子供夫婦がいる世帯
世帯主と子供夫婦と孫がいる世帯
世帯主夫婦がいる世帯
世帯主夫婦と孫がいる世帯
世帯主夫婦と子供がいる世帯
世帯主夫婦と子供と孫がいる世帯
世帯主夫婦と子供がいる世帯
世帯主夫婦と子供と孫がいる世帯
世帯主夫婦と子供夫婦がいる世帯
世帯主夫婦と子供夫婦と孫がいる世帯
世帯主とひとり親がいる世帯
世帯主と孫と親がいる世帯
世帯主と子供とひとり親がいる世帯
世帯主と子供と孫と親がいる世帯
世帯主と子供とひとり親がいる世帯
世帯主と子供と孫と親がいる世帯

96通り
(3 × 2⁵)

表2 新世帯類型の初期段階の設計(抜粋)

プロセス2: 続柄の構成の整理

- 「続柄の構成」を、次のルールで、「新世帯類型用の区分の名称」に変換。
- ① 夫婦でない「子供の配偶者」は、「子供」とした。

表2 新世帯類型の初期段階の設計(抜粋)



- ② 「両親」または「ひとり親」は、4世代以上の世帯については区別せずに「親」に統一した。

プロセス
2②

表2 新世帯類型の初期段階の設計(抜粋)

続柄の構成	新世帯類型用の区分の名称
世帯主と子供の配偶者と孫とひとり親がいる世帯	世帯主と子供と孫と親がいる世帯
世帯主と孫と両親がいる世帯	世帯主と孫と親がいる世帯

「4世代」 = 「4世代に亘る」

【「4世代」の例】 世帯主、孫、両親

世帯主の「子」がないので、続柄は3種類だが、4世代に亘っているので、4世代世帯とする。

プロセス3: 新世帯類型番号の付与

1) ①～⑧のソートキーをExcel関数を用いて作成後、世代内の世帯数の多い順に並び替え。

- ① 「同居世代の数」欄 ⇒ 降順でソート。
- ② 「世帯主は何世代目」欄 ⇒ 昇順でソート。
- ③ 「世帯主夫婦の有無」欄 ⇒ 降順でソート。
- ④ 「続柄の数」欄 ⇒ 降順でソート。
- ⑤ 「子供夫婦の有無」欄 ⇒ 降順でソート。
- ⑥ 「子供の有無」欄 ⇒ 降順でソート。
- ⑦ 「孫の有無」欄 ⇒ 降順でソート。
- ⑧ 「両親の有無」欄 ⇒ 降順でソート。

有無パターン#	直系親族						
	70 祖父母	60 父母	10 世帯主	20 世帯主の配偶者	30 子	40 子の配偶者	50 孫
種類	0 1	0 1 2	1	0 1	0 1	0 1	0 1
80	1	1	1	1	1	1	1
96	1		1	1	1	1	1
76	1	1	1	1	0	1	1
78	1	1	1	1	1	0	1
92	1		1	1	0	1	1
94	1		1	1	1	0	1
64	1	0	1	1	1	1	1
60	1	0	1	1	0	1	1
62	1	0	1	1	1	0	1
74	1	1	1	1	0	0	1
90	1		1	1	0	0	1
58	1	0	1	1	0	0	1
72	1	1	1	0	1	1	1
88	1		1	0	1	1	1
68	1	1	1	0	0	1	1
70	1	1	1	0	1	0	1
84	1		1	0	0	1	1
86	1		1	0	1	0	1
56	1	0	1	0	1	1	1
52	1	0	1	0	0	1	1
54	1	0	1	0	1	0	1
66	1	1	1	0	0	0	1
82	1		1	0	0	0	1
50	1	0	1	0	0	0	1
...

同居世代の数	世帯主は何世代目	世帯主夫婦の有無	続柄の数★	子供夫婦の有無	子供の有無	孫の有無	両親の有無★
降順	昇順	降順	降順	降順	降順	降順	降順
①	②	③	④	⑤	⑥	⑦	⑧
5	3	1	5	1	1	1	0
5	3	1	5	1	1	1	0
5	3	1	5	0	1	1	0
5	3	1	5	0	1	1	0
5	3	1	5	0	1	1	0
5	3	1	5	0	1	1	0
5	3	1	5	0	1	1	0
5	3	1	4	1	1	1	0
5	3	1	4	0	1	1	0
5	3	1	4	0	1	1	0
5	3	1	4	0	0	1	0
5	3	1	4	0	0	1	0
5	3	1	3	0	0	1	0
5	3	0	5	1	1	1	0
5	3	0	5	1	1	1	0
5	3	0	5	0	1	1	0
5	3	0	5	0	1	1	0
5	3	0	5	0	1	1	0
5	3	0	5	0	1	1	0
5	3	0	4	1	1	1	0
5	3	0	4	0	1	1	0
5	3	0	4	0	0	1	0
5	3	0	4	0	0	1	0
5	3	0	3	0	0	1	0
...

プロセス3
1)①~⑧

プロセス3
2)

新世帯類型用の区分の名称

ソート順

新世帯類型番号を付与

世帯主夫婦と子供夫婦と孫と親と祖父母がいる世帯	1
世帯主夫婦と子供夫婦と孫と親と祖父母がいる世帯	1
世帯主夫婦と子供と孫と親と祖父母がいる世帯	2
世帯主夫婦と子供と孫と親と祖父母がいる世帯	2
世帯主夫婦と子供と孫と親と祖父母がいる世帯	2
世帯主夫婦と子供と孫と親と祖父母がいる世帯	2
世帯主夫婦と子供夫婦と孫と祖父母がいる世帯	3
世帯主夫婦と子供と孫と祖父母がいる世帯	4
世帯主夫婦と子供と孫と祖父母がいる世帯	4
世帯主夫婦と孫と親と祖父母がいる世帯	5
世帯主夫婦と孫と親と祖父母がいる世帯	5
世帯主夫婦と孫と祖父母がいる世帯	6
世帯主と子供夫婦と孫と親と祖父母がいる世帯	7
世帯主と子供夫婦と孫と親と祖父母がいる世帯	7
世帯主と子供と孫と親と祖父母がいる世帯	8
世帯主と子供と孫と親と祖父母がいる世帯	8
世帯主と子供と孫と親と祖父母がいる世帯	8
世帯主と子供と孫と親と祖父母がいる世帯	8
世帯主と子供夫婦と孫と祖父母がいる世帯	9
世帯主と子供と孫と祖父母がいる世帯	10
世帯主と子供と孫と祖父母がいる世帯	10
世帯主と孫と親と祖父母がいる世帯	11
世帯主と孫と親と祖父母がいる世帯	11
世帯主と孫と祖父母がいる世帯	12
...	...

表3 新世帯類型の初期段階の設計(抜粋)

プロセス3: 新世帯類型番号の付与

2) 「新世帯類型用の区分の名称」が同じ行に対して、Excel関数を使って同じ連番号を1番から順番に56番まで振り、「小分類符号」を付与。



- ① 親族世帯の新世帯類型を 56 区分に分類
- ② その分類の他に、以下の区分を作成。
 - ・ 親族・非親族同居世帯(1区分)
 - ・ 非親族世帯(1区分)
 - ・ 単独世帯(2区分: 世帯主男性、世帯主女性)
 - ・ 核家族世帯(4区分: 別掲含む)
 - ・ 兄弟姉妹から成る世帯(1区分: 別掲含む)



別掲も含めて 65 区分に分類

Excel上での
一連の作業は
SASプログラムで
処理が可能

プロセス3の①～⑧のキーは全てExcel関数を利用して定義することができた。発表論文に付録として、以下の2つのキーのみ求め方を載せている。

1. 同居世帯数
2. 世帯主は何世代目

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		直系親族												
2	有無パターン#	70	60	10	20	30	40	50	続柄の構成	★同居世代の数④	★エクセル計算③ 同居世代の数	★エクセル計算② 同居世代の数	★エクセル計算① 同居世代の数	から数えて何番目の世代か【注】 世帯主は年配の世代
3		祖父母	父母	世帯主	世帯主の配偶者	子	子の配偶者	孫						
4														
5														
6	種類	0 1	0 1 2	1	0 1	0 1	0 1	0 1						
7	2	0	0	1	0	0	0	1	世帯主と孫がいる世帯	3	1 1	101	101	1
8	18	0	1	1	0	0	0	1	世帯主と孫とひとり親がいる世帯	4	11 1	1101	1101	2
9	82	1	2	1	0	0	0	1	世帯主と孫と両親と祖父母がいる世帯	5	111 1	11101	11101	3
10	83	1	2	1	0	0	1	0	世帯主と子供の配偶者と両親と祖父母がいる世帯	4	1111	11110	11110	3

付録 ソートキーの計算方法

まとめ

- Excel表を使って新世帯類型の96パターンをひとまず特定
- 試行錯誤しながら、エクセルを使って手作業で、パターンの整理・統合や並べ替え
- 完成後に手作業の中味をチェックした結果、全ての手作業はExcel関数を使うことで自動的にできることを確認し、一連のアルゴリズムを整理



次に同様の処理を行うときは、初めの段階からSASでこのアルゴリズムを使い、新世帯類型をSASデータベースに組み込むことができる。



これ以外の分野でも組み合わせの多い分類を作成するとき
に有効に働く。

参考文献

- ① Suoh, S. (2014) *The Construction of a Database System for the Micro Data Analysis of Asian Household Surveys: A Case Study of Indonesia, Income Distribution In Asian Regions: Socioeconomic Analyses Based on Micro Data Since the Second Half of the 20th Century* (edited by Kamiya, D. et al.), Japan Statistical Association, Inc., pp65-117.
- ② 周防節雄(2015)「全国消費実態調査の匿名データからSASのリレーショナルデータベースを完全自動構築するシステムの開発」、統計数理研究所 (http://www.nstac.go.jp/services/pdf/151127_3-2.pdf)
- ③ 伊藤彰彦(2016)『新たな個人消費を創造する新世帯類型別世帯数・世帯人員』、Sinfonica 研究叢書 No.25、(公財) 統計情報研究開発センター
- ④ 新井郁子、米澤香、中川雅義(2016)「新世帯類型による国勢調査マイクロデータの分析結果の概要」、『SASユーザー総会2016論文集』

謝辞

本研究は、独立行政法人日本学術振興会の科学研究費(課題番号26380280:研究代表:伊藤彰彦、「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」)の助成金で行われた。記して謝意に替えたい。