

全国消費実態調査の擬似マイクロデータから 学童・未就学児のいる母子家庭の消費傾向を探る

城元 宏美

関西学院大学 社会学部社会学科
応募部門 Cクラス (学生)

Exploration into the Consumption Tendency by One-Parent Households with Mother and Small Children Using the Pseudo Microdata synthesized from National Survey of Family Income and Expenditure

Hiromi Shiromoto

Kwansei Gakuin University, Faculty of Sociology, Department of Sociology

要旨

両親が揃っている二人親世帯よりも一人親世帯、特に母子家庭は経済的に苦しい立場にあると世間一般では思われることが多い。二人親世帯と一人親世帯を対象にして、収入と支出の様々な指標を比較し、自身の経験から子供の教育費について注目し分析を行った。特に母子家庭では「実支出以外の預貯金への支出がある家庭は、教育費にもお金をかける」という弱い正の相関が存在した。

キーワード:ジニ係数 エンゲル係数 等価可処分所得 相対的貧困率 一人親世帯 母子家庭

1. はじめに

個別指導塾で学生アルバイトをしており、一人親世帯の子供と接する機会がこれまでに多くあった。一人親世帯の子供たちの殆どが母子家庭で、父子家庭の子供は全くといっていいほど在籍していなかった。業務の一環で子供たちの家庭状況を知る機会があったが、どの母子家庭も「金銭的に厳しい」という理由から各季節の特別講習の授業をあまり受講することができないにもかかわらず、保護者からは「必ず公立の学校に合格させてほしい」という要望を受けることがしばしばあった。また、自身の就職活動を通して「地価の高い地域に住む家庭の子供ほど国立や私立の中・高に進学する」という強い相関関係があることを肌で感じ、こうした経験から母子家庭の子供の教育現状と支出状況について関心を持つに至った。

2. 分析について

2.1 分析の方法・定義

本研究は、平成16年全国消費実態調査の匿名データを基に、SAS ユーザー会世話人有志により作成された擬似マイクロデータ(69,131 オブザベーション)を使っている。各オブザベーションには集計用乗率の変数がある。以後、本論文で使う「世帯数」はこの集計用乗率を使った世帯数である¹。

統計局のシングル・マザーの定義では「子と同居している配偶者のいない女性を指し、特に、子が年少と思われる 15～49 歳(再生産年齢)に焦点を当てる」としている(総務省統計研修所研究官室 2006)。なお、シングル・ファーザーの定義もシングル・マザーと同様である。しかし、この定義のシングル・ファーザーとマザーでは二世帯世帯とは限らないし、また公的年金の受給や子供が就業している可能性もあるので、今回の分析を行う際には、「父子・母子世帯」を以下のように定義した。就業人員は 1 人、世帯主の年齢階層は最下層の「24 歳以下」から「45 歳～49 歳」まで、公的年金の受

¹ 集計用乗率を使って復元した総世帯数は、規定問題 1 にあるように、31,887,307 世帯である。

給なし、家族分類は「二世帯(ひとり親)世帯」、「未就学児または学校に通う世帯員がいる」

この定義による母子世帯は 334,631 世帯、父子世帯は 1,0643 世帯であった。

分析手法としては、主に SAS の proc tabulate を用いて視覚的に判断しやすくなる表を作り、proc reg プロシジャを用いて複数の目的変数で従属変数の説明を行うことができる重回帰分析を用いた。

2.2 仮説

自身の経験から以下の仮説を導く。

- ①ひとり親世帯は、二人親世帯と比較すると、年収も子供にかかる教育費も低い。
- ②ひとり親世帯の中では、母子家庭の子供にかかる教育費は父子家庭よりも多い。
- ③母子家庭の世帯主は教養娯楽費、交際を抑え、預貯金への支出を多くし、子供の教育費にお金をかける傾向がある。

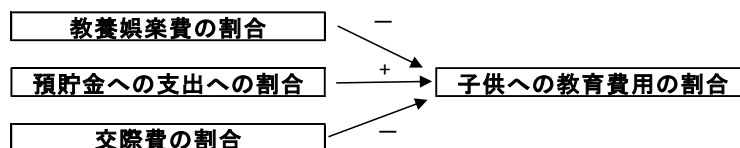


Figure1. 仮説③の媒介予想図

3. 分析の結果

3.1 新擬似マイクロデータ全世帯数に対する等価可処分所得と相対的貧困率

SAS を用いて新擬似マイクロデータ全世帯に対する等価可処分所得を算出し、全世帯における相対的貧困率と貧困線を求めた。なお、相対的貧困率と貧困線を求めるために使用した式は以下の通りである²。

可処分所得 = 実収入 - 非消費支出

等価可処分所得 = 可処分所得 ÷ √世帯人員

貧困線 = 等価可処分所得 ÷ 2

相対的貧困率 = 等価可処分所得が貧困線以下の世帯数 ÷ 全世帯数 × 100

Figure2.のデザインは四方(2016)を参考に作成した。Figure2.に関連する計算処理をした結果、貧困世帯(等価可処分所得が貧困線以下の世帯)は 8,490,670 世帯あり、約 26.6%の世帯が貧困世帯であることが判明した。特に、世帯数累計 25%付近までは急激に世帯数が増加していることから、多くの世帯が貧困線の 78,964 円よりも少ない等価可処分所得で厳しい生活を送っていることが窺える。

また、父子家庭と母子家庭の貧困線から相対的貧困率を求めると、父子家庭では 10,643 世帯中 2,795 世帯が貧困世帯であり、貧困率は 26.3%、母子家庭では 334,641 世帯中 35,077 世帯が貧困世帯であり、貧困率は 10.5%となり、母子家庭よりも父子家庭の世帯の方が貧困率が高いという結果となった。しか

²可処分所得：世帯の「実収入」から支払い義務のある税金や社会保険料などの「非消費支出」を差し引いた後の金額(手取り収入)

等価可処分所得：世帯の可処分所得を世帯員数の平方根で割った値

相対的貧困率：等価可処分所得を高さ順に並べたときに真ん中の所得(中央値)を基準に、その半分に満たない人が占める割合

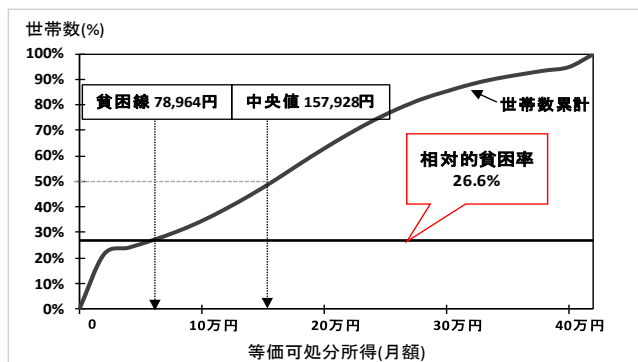


Figure2. 等価可処分所得(月額)と貧困率 (単位:%)

し、父子家庭と母子家庭でサンプル数にやや開きがあるため、誤差が生じている可能性もある。

規定課題2のジニ係数(付録2)からは、家族分類が「単身世帯」、「夫婦のみ世帯」、「二世帯世帯」では世帯主の性別で格差は見られなかった。一方、「二世帯(ひとり親)の世帯」、「三世帯世帯」、「その他の世帯」では「世帯主が女性」の世帯の方が格差の大きいことがわかるが、これは、世帯構造が前者に比べて複雑であることを示唆していると思われる。

3.2 二人親世帯と各ひとり親世帯の年収や様々な支出の比較

家族分類	世帯主の性別	年齢階層	平均値	標準偏差	中央値	最大値	最小値	世帯数
3:二世帯の世帯	1:男	1:20代	430	158	402	1,193	12	476,108
		2:30代	585	225	548	2,851	47	2,810,317
		3:40代	766	303	727	3,539	50	3,139,629
	2:女	1:20代	514	309	603	959	48	4,889
		2:30代	672	497	499	2,515	113	15,557
		3:40代	668	364	538	1,550	69	19,503
	全世帯	1:20代	431	161	402	1,193	12	480,998
		2:30代	586	227	548	2,851	47	2,825,874
		3:40代	765	304	727	3,539	50	3,159,132
4:二世帯(ひとり親)の世帯	1:男	2:30代	440	206	470	778	153	2,623
		3:40代	729	582	514	2,574	252	8,018
	2:女	1:20代	187	84	181	535	64	16,804
		2:30代	342	239	269	1,621	21	142,048
		3:40代	496	334	397	2,067	44	175,789
	全世帯	1:20代	187	84	181	535	64	16,804
		2:30代	343	238	271	1,621	21	144,672
		3:40代	506	352	408	2,574	44	183,808

Table1. 年間収入 (単位:万円)

Table1.に、世帯の年間収入(万円)について、家族分類別、世帯主性別、世帯主の年齢階層別の基本統計量を示す。本表を算出する際、あらかじめ作成した父子・母子家庭と比較する対象世帯として、家族分類の二世帯の世帯、つまり両親が揃っている二人親世帯についても、次の様に定義を行った。定義は以下の通りである。就業人員は0~2人、世帯主の年齢階層は最下層の「24歳以下」から「45歳~49歳」まで、公的年金の受給なし、家族分類は「二世帯世帯」、「未就学児または学校に通う世帯員がいる」

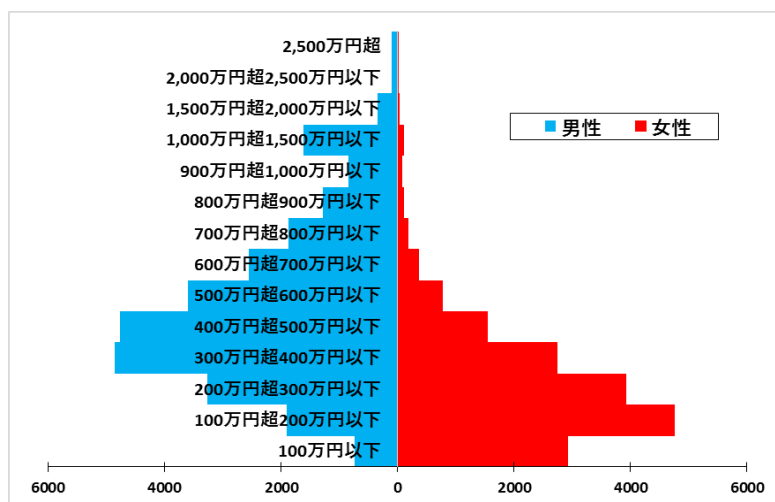


Figure.3 平成19年 平均年収 (単位:千人)

の20代の年間収入の少なさは顕著である。Figure3.は民間給与実態統計調査—調査結果報告(国税庁長官官房企画課2011)のデータから筆者がExcelで作成した平成19年の男女別の平均

二世帯世帯、「未就学児または学校に通う世帯員がいる」

この定義による世帯主が男性の二世帯世帯は6,426,054世帯、世帯主が女性の二世帯世帯は39,949世帯、二世帯世帯全体は6,466,003世帯であった。

Table1.から世帯主の性別年齢階層にかかわらず、平均値・中央値共にひとり親の世帯の年間収入は二世帯世帯よりも低い。特に母子家庭

年収の分布のグラフである。グラフから見て取れるように男性よりも女性の方が年収が著しく低いことがわかる。新擬似マイクロデータはこの結果と矛盾しない。

家族分類	世帯主の性別	年齢階層	平均値	標準偏差	中央値	最大値	最小値	世帯数
3:二世帯の世帯	1:男	1:20代	6,667	12,861	0	175,920	0	476,108
		2:30代	17,278	24,551	11,252	559,042	0	2,810,317
		3:40代	39,624	51,927	23,696	661,644	0	3,139,629
	2:女	1:20代	19,016	16,372	20,337	47,260	0	4,889
		2:30代	21,600	25,363	12,846	92,227	0	15,557
		3:40代	40,948	52,374	29,993	227,091	0	19,503
	全世帯	1:20代	6,792	12,961	0	175,920	0	480,998
		2:30代	17,302	24,558	11,252	559,042	0	2,825,874
		3:40代	39,632	51,929	23,698	661,644	0	3,159,132
4:二世帯(ひとり親)の世帯	1:男	2:30代	43,793	69,596	5,265	178,638	0	2,623
		3:40代	18,368	41,985	3,181	207,731	0	8,018
		2:30代	12,190	17,066	5,602	90,526	0	142,048
	2:女	1:20代	3,413	6,551	1,542	45,780	0	16,804
		2:30代	12,190	17,066	5,602	90,526	0	142,048
		3:40代	43,794	63,916	20,930	378,331	0	175,789
	全世帯	1:20代	3,413	6,551	1,542	45,780	0	16,804
		2:30代	12,763	19,787	5,538	178,638	0	144,672
		3:40代	42,684	63,331	19,241	378,331	0	183,808

Table2. 教育費 (単位:円)

Table2.は Table1.と同様に、proc tabulate を用いて新擬似マイクロデータから月額教育費を算出した。教育費は、ひとり親世帯よりも二世帯世帯の方が金額が高い傾向があるが、ひとり親世帯だけに注目すると、世帯主が30・40代では中央値を見ると父子家庭よりも母子家庭の方が教育費が高いことがわかる。

3.3 教育費についてのさらなる分析

教育費については各世帯の収入によって負担の比重が変化する。そこで、世帯主の性別のひとり親世帯の収入総額に対する教育費の割合について対応のないt検定を行った。p値は0.59となり、有意差はないという結果となった。同じく年収と月額収入総額でt検定を行ったところ、年収は $p < .0001$ 、月額収入総額は $p = 0.028$ となり、どちらも性別による差があることが判明した。

家族分類	世帯主の性別	年齢階層	平均値	標準偏差	中央値	最大値	最小値	世帯数
4:二世帯(ひとり親)の世帯	1:男	2:30代	39.5	53.5	2.8	129.2	0.0	2,203
		3:40代	11.3	18.8	7.4	83.7	0.0	5,645
	2:女	1:20代	3.2	5.3	1.7	36.9	0.0	16,804
		2:30代	9.7	28.7	5.5	459.2	-169.6	137,105
		3:40代	29.6	41.5	15.1	232.5	-121.8	163,989
	全世帯	1:20代	3.2	5.3	1.7	36.9	0.0	16,804
		2:30代	10.1	29.5	5.5	459.2	-169.6	139,308
		3:40代	29.0	41.1	14.3	232.5	-121.8	169,634

Table4. 月当りの等価可処分所得に対する教育費の割合 (単位:%)

また、Table4.より「母子家庭の方が父子家庭よりも月額収入総額に対する教育費の割合が、平均値、中央値共に高い」ことが読み取れるので、母子家庭の方が父子家庭よりも子供の教育に力を入れている傾向があると言える。更に、相関分析を行ったところ、「等価可処分所得に対する預貯金への支出の割合が高い世帯は、等価可処分所得に対する教育費の割合も高い」という弱い正の相関($r=0.294$)が存在した。この結果から、「預貯金への支出の割合が高いことは、子供の教育に関心がある」と言える。子供の教育に金銭を多く使うということは、世帯主自身の自由なお金、旅行などが含まれる「教養娯楽費」や「交際費」を節約して預貯金を行っているのではないかと

仮説を立て、「月あたりの等価可処分所得に対する教育費の割合」を従属変数とし、「月あたりの等価可処分所得に対する教養娯楽費の割合」、「月あたりの等価可処分所得に対する預貯金への支出の割合」、「月あたりの等価可処分所得に対する交際費の割合」を独立変数とする重回帰分析のモデルを作成し、分析した。

変数名	係数	標準誤差	p値
教養娯楽費の割合	0.241	0.047	<.0001 **
預貯金への支出の割合	0.060	0.007	<.0001 **
交際費の割合	0.371	0.059	<.0001 **
切片	5.105	1.480	
N	1153obs		
adjR ²	0.144 **		

**p <.01, *p <.05, +p <.10

Table5. 教養娯楽費と交際費と預貯金への支出の教育費への重回帰分析

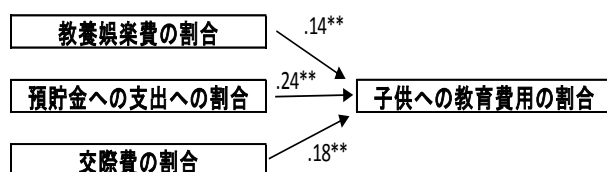


Figure4. 教養娯楽費と交際費と預貯金への支出の教育費への重回帰分析のモデル図

左の Table5.より、母子家庭では、「月当たりの等価可処分所得に対する教養娯楽費の割合」、「月当たりの等価可処分所得に対する預貯金への支出の割合」、「月当たりの等価可処分所得に対する交際費の割合」の3つの変数で教育費の約14.4%を予測できることがわかった。(自由度調整済み決定係数<adjR²=0.144>より)

各変数の効果を確認したところ、教養娯楽費の割合が高いほど教育費の割合が高く(b=.241 p<.0001)、預貯金への支出の割合が高いほど教育費も高く(b=.060 p<.0001)、交際費の割合も高いほど教育費も高く(b=.371 p<.0001)、という結果が得られ、この重回帰分析のモデルは1%水準で有意であった。

左の Figure4.は重回帰分析の係数を標準化した推定値で説明をした重回帰分析のモデル図である。

4. 結論・まとめ

4.1 結論:仮説について

以上の分析から以下のことが判明した。

- ①ひとり親世帯は、二人親世帯と比較すると、年収も子供にかかる教育費も低い。
- ②ひとり親世帯の中では、母子家庭の子供にかかる教育費は父子家庭よりも多い。
- ③母子家庭の世帯主は教養娯楽費、交際費、預貯金への支出が多いが、子供の教育費にお金をかける傾向がある。

仮説①、②に関して仮説は支持されたが、仮説③に関しては棄却された。

4.2 まとめ

本稿では「2004年(平成16年)全国消費実態調査」の匿名データを基に作成された擬似マイクロデータを用いて様々な分析を行った。年収については国税庁のデータと矛盾しない結果が得られ、男性と女性間に有意な差があることが判明した。その理由として、パートタイマーやアルバイトなどの非正規雇用が多いことと、正規雇用の社員であっても男性より女性は昇進のスピードが遅いということが考えられる。女性の結婚や出産、育児などのライフ・イベントが年収の差に影響しているのではないだろうか。また、父子家庭よりも母子家庭の方が教育費で、金額でも収入に対する割合でも高いという分析結果は、「母子家庭の方が子供に教育に関心がある」ということを示してい

ると考えられる。そのため重回帰分析で「月当りの等価可処分所得に対する預貯金への支出の割合が高い」ということは将来の進学のために貯金をしていると予想できるが、教養娯楽費と交際費が教育費に正の影響を与えていることを説明することは難しい。これはデモグラフィックデータを用いて分析ができれば、相関関係などをより一層正確に説明することができるようになると思われる。

参考文献

- [1] 総務省統計研修所研究官室(2006)「シングル・マザーの最近の状況 その3」、『統計』、財団法人日本統計協会発行、平成 18 年 11 月号
- [2] 総務省統計研修所研究官室(2007)「シングル・ファーザーの最近の状況」、『ESTRELA』、財団法人統計情報研究開発センター、平成 19 年 3 月号
- [3] 四方克成 (2016)「貧困・中流・富裕世帯における教育費の実態調査」、『SAS ユーザー総会 2016 論文集』、pp.365-379
- [4] 国税庁 (2011)「平成 23 年分 民間給与実態統計調査 一調査結果報告一」、国税庁 長官官房 企画課

付録 3 自由課題 SAS プログラム

```
/*母子家庭の重回帰分析(自由課題)*/ options nocenter;
%let drive=D;
libname giji "&drive:%データコンペ2018%giji";
libname engel "&drive:%データコンペ2018%sasds";

%include "&drive:%データコンペ2018%saskitei%kiteiFormat.sas";

data MOMandChild;
  set engel.spending;
  label engel="エンゲル係数"
        DSPincome="可処分所得"
        EQincome="等価可処分所得"
        YRincome="年間収入(単位万円)"
        EQLMpercent="等価可処分所得に対する教養娯楽代の割合"
        EQSOCpercent="等価可処分所得に対する交際費の割合"
        EQEDUpercent="等価可処分所得に対する教育費の割合"
        EQSRGpercent="等価可処分所得に対する預貯金支出の割合";

  keep sex X04 X07 X09 HHkind X12 X13 YRincome Y002 Y003 Y018
        ALLspending Y083 Y137 Y141 Y159 Y179 Y191
        engel DSPincome xEQincome EQincome weight Ageclass
        EQLMpercent EQEDUpercent EQSOCpercent EQSRGpercent;

  DSPincome=Y003-Y179; *可処分所得=実収入-非消費支出;
  xEQincome=DSPincome/SQRT(X03); *等価可処分所得=可処分所得÷√世帯人員数;
  EQincome=int(xEQincome); *等価可処分所得(円未満四捨五入);

  EQLMpercent=Y141/EQincome*100; *等価可処分所得に対する教養娯楽代の割合;
  EQSOCpercent=Y159/EQincome*100; *等価可処分所得に対する交際費の割合;
  EQEDUpercent=Y137/EQincome*100; *等価可処分所得に対する教育費の割合;
  EQSRGpercent=Y191/EQincome*100; *等価可処分所得に対する預貯金支出の割合;

  select;
    when (X09=<6) Ageclass=1; *20代;
    when (X09=<8) Ageclass=2; *30代;
    when (X09=<10) Ageclass=3; *40代;
    otherwise ;
  end;

  if (X04=1) AND (X12=2 OR X13=2) AND (X09=<10)
    AND (HHkind=4) AND (sex=2) AND (Y018=0);

run;

proc reg data=MOMandChild;
  weight weight;
  model EQEDUpercent = EQLMpercent EQSOCpercent EQSRGpercent/ selection=stepwise vif;
  output out=P2 P=predict R=resid;
  print stb r;
  title "母子家庭重回帰分析";
run; quit;
```

付録1 規定課題1用 SAS プログラム

```

/* kitei1.sas */ options nocenter;

libname kitei "D:¥データコンペ2018¥giji";
%include "D:¥データコンペ2018¥saskitei¥kiteiformat.sas";

proc tabulate data=kitei.zensho2004gijimicro vardef=WDF format=comma10.0;
  class HHkind sex;
  var YRincome / weight=weight;
  table (HHkind ALL="全世帯"),
        (sex ALL="全世帯")*(YRincome="平均値"*MEAN="")
        (sex ALL="全世帯")*(YRincome="標準偏差"*STDdev="")
        (sex ALL="全世帯")*(YRincome="世帯数"*SUMWGT="");
  format sex sex.;
  format HHkind HHkind.;
  title "規定問題1：解答";
run;

```

付録1 規定課題1 図表

規定問題1:解答									
家族分類	世帯主の性別		全世帯	世帯主の性別		全世帯	世帯主の性別		全世帯
	1:男	2:女		1:男	2:女		1:男	2:女	
	平均値	平均値	平均値	標準偏差	標準偏差	標準偏差	世帯数	世帯数	世帯数
1:単身世帯	416	266	338	267	165	232	3,187,182	3,432,001	6,619,183
2:夫婦のみの世帯	557	539	557	361	331	360	7,600,840	143,504	7,744,344
3:二世帯の世帯	733	686	732	373	366	373	11,709,238	147,052	11,856,290
4:二世帯(ひとり親)の世帯	689	459	558	402	326	378	1,003,841	1,324,421	2,328,262
5:三世帯の世帯	923	677	904	459	421	461	2,545,855	210,980	2,756,835
6:その他の世帯	778	576	724	456	429	458	428,659	153,734	582,393
全世帯	662	356	610	396	281	396	26,475,615	5,411,692	31,887,307

【規定課題1のプログラム解説】

proc tabulate では、集計用乗率の変数 weight を使う場合は、注意を要する。使う際のポイントは、以下の二つのオプションである。

① vardef=WDF

集計用乗率を考慮した標準偏差を求めるには、このオプションが必要になる。

② SUMWGT=

通常、件数を求めるときは、予約語「N」を使うが、N には weight 文が効かない。集計用乗率を件数に反映させるには SUMWGT=を使う。

付録2 規定課題2用 SAS プログラム

プログラム①

```
/* divide_DS.sas */
libname giji "D:¥擬似マイクロデータの保存されている完全パス";
%macro print(dsname);
proc print data=&dsname (obs=10); title "data=&dsname"; run;
%mend;

data original; set giji.zensho2004gijimicro; *元の擬似マイクロデータセット;
keep Y001 X08 X11 weight;
rename Y001=income X08=sex X11=HHkind;
run; %print(original);

*★元のデータセット(全世帯)を「20分割+全世帯=21分類」;
*家族分類(6分類+全家族分類=7分類)×世帯主性別(2分類+男女=3分類)=21分類;
*以下の順に分割していく。
  ① HH1all HH2all HH3all HH4all HH5all HH6all
  ② HH1M HH1F HH2M HH2F HH3M HH3F HH4M HH4F HH5M HH5F HH6M HH6F
  ③ HHallM HHallF HHallMF;

*① HH1all HH2all HH3all HH4all HH5all HH6all;
data HH1all HH2all HH3all HH4all HH5all HH6all;
set original;
select (HHkind);
  when(1) output HH1all;
  when(2) output HH2all;
  when(3) output HH3all;
  when(4) output HH4all;
  when(5) output HH5all;
  when(6) output HH6all;
  otherwise put "★ HHkind ? " HHkind=;
end; *変数 : sex HHkind income weight;
run;

*② HH1M HH1F HH2M HH2F HH3M HH3F HH4M HH4F HH5M HH5F HH6M HH6F;
%macro wakeru_sex(No);
data HH&No.M HH&No.F; set HH&No.all;
select (sex);
  when(1) output HH&No.M;
  when(2) output HH&No.F;
  otherwise put "★ sex ? " sex=;
end; *変数 : sex HHkind income weight;
run;
%mend;

%wakeru_sex(1);
%wakeru_sex(2);
%wakeru_sex(3);
%wakeru_sex(4);
%wakeru_sex(5);
%wakeru_sex(6);

*③ HHallM HHallF;
data HHallM HHallF; set original;
select (sex);
  when(1) output HHallM;
  when(2) output HHallF;
  otherwise put "★ sex ? " sex=;
end; *変数 : sex HHkind income weight;
run;

data HHallMF; set original; run;
```


プログラム②

```
/* gini係数計算プログラムmacro化.sas */ options nocenter mtrace macrogen;
*マクロ変数名
  ds : データセット名
  var : ジニ係数を計算したい変数名
  weight : 集計用乗率の変数名
  digit_below_zero : 集計用乗率の小数点以下第何位(0,1,2・・・)まで使う
  NOPRINT:計算の途中経過を結果ビュー画面に表示する(半角ブランク)か否か(*)
;
%macro gini(ds,var,weight,digit_below_zero,NOPRINT);
data original; set &ds;
proc sort data=original; by &var; run; *ジニ係数を計算する変数を昇順にソートする;

proc means data=original NOPRINT; output out=minWGT; run;

data _NULL_; set minWGT; if _STAT_="MIN" then do; minWGT=&weight;
  call symput("minWGT",&weight); end; run;

%put &minWGT;

data original1; set original; newWGT=&weight/&minWGT; run; &NOPRINT%print(original1);

data original2; drop &weight TTLobs;
  set original1 end=owari; repeat=round(newWGT*10**&digit_below_zero);
  TTLobs+repeat; if owari then call symput('TTLobs',TTLobs); *★膨らませるobservationの数;
run; &NOPRINT%print(original2);

%put &TTLobs;

data step0; drop i repeat;
  set original2; do i=1 to repeat; weight=1; output; end; run;

data step1; set step0 end=owari;
  keep ACCweight ACC&var;
  ACCweight+weight;
  ACC&var+&var;
  if owari;
run; &NOPRINT%print(step1);

data step2; merge step0 step1; run; &NOPRINT%print(step2);

data step3; set step2;
  retain xACCweight xACC&var;
  if _N_=1 then do; xACCweight=ACCweight; xACC&var=ACC&var; end;
run; &NOPRINT%print(step3);

data step4; set step3 end=owari;
  standardWGT =weight/xACCweight; ACCstnWGT+standardWGT;
  standard&var=&var/xACC&var; ACCstn&var+standard&var;
  if owari then put "★検算★" ACCstnWGT= ACCstn&var=;
run; &NOPRINT%print(step4);
```

```

data onlyACCstn&var; set step4; keep ACCstn&var standardWGT;
  rename ACCstn&var=katei;
run; &NOPRINT%print(onlyACCstn&var);

data step5;
  set onlyACCstn&var end=owari; if owari then delete;
run; &NOPRINT%print(step5);

data step6; katei=0; run; &NOPRINT%print(step6);

data step7; set step6 step5; jotei=katei; keep jotei; run; &NOPRINT%print(step7);

data step8; merge onlyACCstn&var step7; run; &NOPRINT%print(step8);

data step9; set step8 end=owari;
  daikei_menseki=(jotei+katei)*standardWGT/2;
  TTLdaikei_menseki+daikei_menseki;
  TTLstandardWGT+standardWGT; *★検証 : weight変数の累計;
run; &NOPRINT%print(step9);

data step10; keep dataset gini;
  set step9 end=owari;
  length dataset $ 8;
  if owari then do arch_menseki=0.5-TTLdaikei_menseki;
    gini=arch_menseki/0.5;
    dataset="&ds";
    output;
  end;
run;

proc print data=step10; title; run;
%mend gini;

%macro print(dsname); *★proc print用;
proc print data=&dsname (obs=10); title "&ds.: &dsname"; run;
%mend;

```

プログラム③

```
/* gini係数マクロ実行.sas */
*★元のデータセット(全世帯)を「20分割+全世帯=21分類」;
*家族分類(6分類+全家族分類=7分類)×世帯主性別(2分類+男女=3分類)=21分類
  ① HH1all HH2all HH3all HH4all HH5all HH6all
  ② HH1M HH1F HH2M HH2F HH3M HH3F HH4M HH4F HH5M HH5F HH6M HH6F
  ③ HHallM HHallF HHallMF;

%macro dscopy(newDSname); data x&newDSname; set step10; run; %mend dscopy;

%gini(HH1M,income,weight,1,*) %dscopy(HH1M)
%gini(HH2M,income,weight,1,*) %dscopy(HH2M)
%gini(HH3M,income,weight,1,*) %dscopy(HH3M)
%gini(HH4M,income,weight,1,*) %dscopy(HH4M)
%gini(HH5M,income,weight,1,*) %dscopy(HH5M)
%gini(HH6M,income,weight,1,*) %dscopy(HH6M)
%gini(HHallM,income,weight,1,*) %dscopy(HHallM)

%gini(HH1F,income,weight,1,*) %dscopy(HH1F)
%gini(HH2F,income,weight,1,*) %dscopy(HH2F)
%gini(HH3F,income,weight,1,*) %dscopy(HH3F)
%gini(HH4F,income,weight,1,*) %dscopy(HH4F)
%gini(HH5F,income,weight,1,*) %dscopy(HH5F)
%gini(HH6F,income,weight,1,*) %dscopy(HH6F)
%gini(HHallF,income,weight,1,*) %dscopy(HHallF)

%gini(HH1all,income,weight,1,*) %dscopy(HH1all)
%gini(HH2all,income,weight,1,*) %dscopy(HH2all)
%gini(HH3all,income,weight,1,*) %dscopy(HH3all)
%gini(HH4all,income,weight,1,*) %dscopy(HH4all)
%gini(HH5all,income,weight,1,*) %dscopy(HH5all)
%gini(HH6all,income,weight,1,*) %dscopy(HH6all)
%gini(HHallMF,income,weight,1,*) %dscopy(HHallMF)

*21個のジニ係数を規定問題2の表様式に整理する;
data col1; set xHH1M xHH2M xHH3M xHH4M xHH5M xHH6M xHHallM; rename gini=giniM; run;

data col2; set xHH1F xHH2F xHH3F xHH4F xHH5F xHH6F xHHallF; rename gini=giniF; run;

data col3; set xHH1all xHH2all xHH3all xHH4all xHH5all xHH6all xHHallMF;
  rename gini=giniMF;
run;

data gini; merge col1 col2 col3;
label dataset="家族分類" giniM="世帯主:男" giniF="世帯主:女" giniMF="全世帯";
run;

proc format;
  value $dataset
  HH1all="1.単身世帯"
  HH2all="2.夫婦のみ世帯"
  HH3all="3.二世代世帯"
  HH4all="4.二世代(ひとり親)世帯"
  HH5all="5.三世代世帯"
  HH6all="6.その他の世帯"
  HHallMF="全世帯";
run;
```

```
proc print data=gini label NOOBS;
  format dataset $dataset.;
  format giniM F5.2;
  format giniF F5.2;
  format giniMF F5.2;
  title "規定課題2：年間収入のジニ係数"; run;
```

【規定問題2のプログラム解説】

ジニ係数を求める規定問題2では、以下の三つのプログラムから成り、この順番に実行すれば、上に示す様に、規定問題2の解答が自動的に得られる。

- ① divide_DS.sas
- ② gini 係数計算プログラム macro 化.sas
- ③ gini 係数マクロ実行.sas

プログラム①では、最終的に全部で 21 個のジニ係数を求めるために、元の擬似マイクロデータを 21 個の分類ごとに 21 個のデータセットに分割した。その際、プログラム②、③で SAS マクロを容易に使える様に、データセット名に工夫を凝らした。

プログラム②にあるマクロ **gini** では、

(1)データセット名、(2)変数名、(3)集計用乗率の変数名、(4)集計用乗率の小数点何位まで使う、(5)計算の途中経過を表示するか否かの五項目を指定する。ジニ係数を算出する際は、集計用乗率を使って、あらかじめ各オブザベーションを複数コピーしている。

このマクロは汎用的に使用できる様に配慮した。もし、集計用乗率の変数がないデータセットでジニ係数を計算する場合は、事前に全てのオブザベーションに **weight=1** を追加しておけばよい。

プログラム③では、マクロ **gini** を実行して、21 個のジニ係数を計算し、規定問題の求める表様式に自動的に並べている。実行には約 4 分少々かかった。

付録 2 規定課題 2 図表

規定課題2:年間収入のジニ係数			
家族分類	世帯主:男	世帯主:女	全世帯
1:単身世帯	0.31	0.32	0.34
2:夫婦のみ世帯	0.31	0.31	0.31
3:二世帯世帯	0.26	0.28	0.26
4:二世帯(ひとり親)世帯	0.30	0.36	0.35
5:三世帯世帯	0.26	0.34	0.27
6:その他の世帯	0.30	0.37	0.33
7:全世帯	0.31	0.38	0.33