

マイクロデータ分析のための演習用教材の作成方法

～一般用マイクロデータ詳細品目版及び擬似マイクロデータによる事例～

河野 真理子[†]

和田 かず美[†]

Preparation of Teaching Materials for Microdata Analysis: Case Examples Using Official Pseudo Microdata

KAWANO Mariko

WADA Kazumi

本稿では、統計解析システム R の基礎的な利用方法を教えるマイクロデータ分析の講座の演習資料を作成した経験に基づき、教材作成に必要なデータの分析手法を、実際の分析事例とともに紹介する。さらに、このような教育目的で使用するデータの要件についても考察する。マイクロデータ分析の演習に利用可能な公的データには、独立行政法人統計センターが提供する一般用マイクロデータを使った。

キーワード：公的統計、擬似マイクロデータ、統計教育、マイクロデータ分析

Based on our experiences of lecturing on microdata analysis in R language, the statistical analysis system, we propose some of the data analyzing methods that should be contained in teaching materials together with basic R codes. We further discuss the basic requirements of microdata for such an educational purpose. The microdata we used in this paper is a new pseudo microdata that was developed recently using the Anonymous Microdata of *the National Survey of Family Income and Expenditure*, and uploaded in the internet by the National Statistics Center (NSTAC).

Key words: Official Statistics, Synthetic Microdata, Statistical Education, Microdata Analysis

[†] 独立行政法人統計センター統計情報・技術部統計技術研究課

1. はじめに

本稿は、統計解析システム R による初心者向けのマイクロデータ分析の演習教材（和田・河野，2017）を作成した経験に基づき、教材作成の基本的な手順と有用な手法について紹介する。演習教材作成時に使用したデータは、独立行政法人統計センター（以下、統計センター）が公開していた「教育用擬似マイクロデータ」と、現在提供されている「一般用マイクロデータ詳細品目版」を用いた。併せて、「一般用マイクロデータ詳細品目版」を用いた分析事例を紹介する。また、マイクロデータ分析演習に用いる教材として、どのような点が重要であるかについてもデータのユーザーの観点から考察する。

まず、第2節では、マイクロデータ分析の演習に利用可能な公的データにはどのようなものがあるのかを概観し、使用した「擬似マイクロデータ」及び「一般用マイクロデータ」について紹介する。第3節は、初歩的なマイクロデータ分析の方法論について述べ、併せて「一般用マイクロデータ詳細品目版」による分析事例を紹介する。第4節は、「一般用マイクロデータ詳細品目版」及び「擬似マイクロデータ」を用いた分析経験から、両者の相違点についてまとめ、それが分析にどのような影響があるかを事例紹介する。最後に第5節で、この二種類の擬似マイクロデータの分析結果を元に、マイクロデータ分析に用いることのできるデータの要件についてまとめる。なお、付録1に本稿で作成した図の R のコード、付録2に一般用マイクロデータの符号表の一部、付録3及び付録4に擬似マイクロデータの符号表を添付した。さらに実際に作成した演習教材については、和田・河野（2017）として公開している。本文では、「擬似マイクロデータ」を擬似マイクロ、「一般用マイクロデータ」を一般マイクロと呼ぶ。

2. 教育用に利用できる公的なマイクロデータについて

公的統計は、従来国の政策決定の基礎資料と位置づけられ、その作成及び提供に関して基本的な事項を定める昭和22年成立の旧統計法に基づき、主に統計表の形で公表され、マイクロデータの利用は新たな統計の作成や公的な学術研究等の目的に例外的に認められるのみであった。この旧統計法が平成19年5月に抜本改正され、公的統計を社会の情報基盤とする新統計法（平成19年法律第53号）が可決・成立、平成21年4月から全面施行された。新統計法には、国民の財産である公的統計のさらなる利用を図るため、新たな二次的利用制度の一つとして、匿名データの作成及び提供について明記されている。匿名データとは、統計調査のマイクロデータを、調査客体が特定されないよう、匿名化の処理を施して提供するもので、受益者負担を原則とする複製費用相当ではあるが有償で、利用目的は学術研究等の目的に限定される。加えて、利用者の範囲や利用場所、環境、保管等についても規定があり、大学生の基礎教育レベルあるいは社会人向けの統計教育等の演習に利用するにはハードルが高い。

このような教育目的に対応するため、より利用上の制約が少ない公的なマイクロデータを提供する試みとして、統計センターにおいて、平成16年全国消費実態調査に基づく擬似データである「教育用擬似マイクロデータ」が開発され、平成23年8月に試行提供開始、平成24年8月に利用範囲が教育用以外にも拡大され、名称も「擬似マイクロデータ」となった。これは、メールで申請者の氏名所属連絡先に加えて、利用目的と利用者の範囲を明記した申出により、誰でも無償で利用が可能であったため、総務省統計研究研修所（当時の名称は統計研修所）における統計解析システム R の講座の演習教材として採用した。

ところが、その後平成26年を始期とする「公的統計の整備に関する基本的な計画」（第Ⅱ期基本計画）において、「一般用マイクロデータ」の提供推進が謳われ、その作成方法に関する検討が行われた後、平成28年3月に「一般用マイクロデータ」の提供開始に伴い、「擬似マイクロデータ」の試行提供が終了した。その経緯は2.2節で解説する。

「一般用マイクロデータ」は、大学・高等学校等における統計教育やデータサイエンスの学習用データ、システム検証用データ等への利活用を目的として、平成21年全国消費実態調査に基づき作成された新しい擬似マイクロデータである。誰でも自由に無償でインターネットからダウンロードすることができ、氏名などの登録は必要としないという点で、「擬似マイクロデータ」よりも利用しやすさが向上している。一方で、2.2節で詳しく述べるが、当初公開されたものはマイクロデータ分析演習への利用は困難であった。その後、利用者の意見を反映して改善され、より大規模な詳細品目版が平成28年に公開されたため、統計研究研修所における教材も「擬似マイクロデータ」から「一般用マイクロデータ」に移行した。

「擬似マイクロデータ」も「一般用マイクロデータ」も、乱数による擬似データであるため、分析結果を実証研究に使用することはできないが、実際の統計調査データではないという点で匿名データと異なっている。本節以下では、その作成方法の違いを中心に解説する。

2.1 「擬似マイクロデータ」について

「擬似マイクロデータ」（以下、擬似マイクロ）は、統計センターが独自に開発した、二人以上の勤労者世帯を対象にしたレコード数32,027のデータで、世帯主の性別、年齢階級別などの世帯属性についての質的変数（カテゴリ変数）が14項目、量的変数は、支出項目149項目、収入項目が34項目収録された。世帯属性は、性別、年齢、就業・非就業の別といった世帯主に関する調査事項と世帯区分、世帯人員階級といった世帯に関する調査事項を含む。

山口・伊藤・秋山（2013）によれば、擬似マイクロは、調査票情報から高次元の集計表を作成し、その高次元の集計表から調査票情報に近似したマイクロデータを作成している。個別データではなく集計結果から作成するため、擬似的なマイクロデータといえるが、調査票情報である個票データの分布にできる限り近似するよう、集計表の各セルの量的属性値の統計量（平均、分散及び共分散）に基づく多変量正規乱数を発生させ、擬似的な値が生成されていた。このため、擬似マイクロの分布特性は、調査票情報のそれに近似し、分析を行った場合、基となったものに近い集計表が復元される。高次の統計量（歪度、尖度等）が再現されることは保証されないが、データが正規分布に近い場合には、統計の量的属性の層ごとの平均値、標準偏差と、相関はほぼ再現されていた。

2.2 「一般用マイクロデータ」について

擬似マイクロの後継となった「一般用マイクロデータ」（以下、一般マイクロ）は、総務省統計局の委託を受け、統計センターが作成・提供している。統計表に基づき、擬似乱数により作成された擬似データであるという点について、一般マイクロと擬似マイクロは同じであり、調査票情報であるマイクロデータから作成するのではなく、集計した統計表に基づいて作成することにより、匿名データと比較して利用上の制約の大幅な緩和を図っている。しかし、統計センターが作成し試行提供した擬似マイクロ作成用の高次元の集計表は詳細すぎて、秘匿性の観点から公表は困難であったため、一般マイクロは公表が可能なレベルでの基本的な統計量（度数及び数量（平均、標準偏差等））に関する統計表を統計局が作成・公表した後に、統計センターがその統計表に基づいて一般マイクロを作成・公表することとなった。

最初に開発された一般マイクロは、平成21年全国消費実態調査に基づき、質的属性は3大都市圏か否か、世帯主の年齢階級、世帯人員、住宅所有関係、高齢者世帯か否かの6項目と量的属性は年間収入、消費支出及び十大費目の12項目が収録され、変数の数が少なくまた各変数の度数及び数量は再現をしているが、変数間の相関は再現されていないという、擬似マイクロと異なる大きな特徴があった。つまり、擬似マイクロが平均値ベクトルと分散共分散行列の情報に基づいて作

成されていたのに対し、この一般マイクロは平均値ベクトルと個々の変数の分散のみに基づき再現されていたために、変数間の相関について元となったデータセットの情報が再現されておらず、統計表を作成する演習には問題がないが、個々のデータの変数間の関係性に着目するマイクロデータ分析に使用するデータとしては適さない。このような利用者の意見が反映された結果、平成28年12月22日に量的変数間の相関構造も考慮をした改訂版が提供された。この一般マイクロの改訂版には、質的変数が7項目と復元ウェイト、量的変数が11項目収録されている。さらに、平成29年6月23日には新たに詳細品目版の提供が開始された。この詳細品目版は、相関構造も再現する改訂版をもとに、410品目分類のデータを加えて作成されている。

一般マイクロの詳細品目版は、まず世帯属性別に十大費目について平均値ベクトルと分散共分散行列の情報を再現する対数正規乱数データが作成されている。次に、410品目分類については、世帯属性別の特化係数を反映した品目分類別の構成比を按分比率として、十大費目に乗算して得られた按分値をもとにした乱数により作成されている¹。

一般マイクロの詳細品目版のデータ構成は、世帯属性等が7項目、集計用乗率及び収支項目が13項目となっている。質的属性には年齢階級、就業・非就業の別、就業人員といった世帯主に関する調査事項と3大都市圏か否か、世帯人員、住居所有関係の世帯に関する調査事項に区別されている。量的属性については、収入項目は年間収入のみであるが、支出項目については11項目あり、さらには410項目の詳細品目別支出が収録されている。

3. 教材作成のための分析法

通常、業務として行う一般的なデータ分析では、事前に分析目的が存在する。使用すべきデータも決まっている場合が多く、データと目的が決まれば、分析手法もそこから決まる。一方で、分析目的が教材作成である場合、最初に利用できるデータを探し、次にそのデータに合う分析手法を探す。いくつかの分析手法により試行錯誤した後、明らかな結果が出るものを選び、できるだけ教材の利用者の興味を惹くような分析シナリオを作成していく。そのために、どのような教材が作成できるかは、データに大きく依存することになる。本節では、一般マイクロの詳細品目版を例に、分析の手順とそのためにより有用な手法について紹介する。一般マイクロは、質的変数と量的変数から構成され、データとともにデータレイアウトと符号表が提供されている。この符号表にある世帯属性等と十大費目部分について、付録2に示すが、一般用マイクロデータの利用サイト (<http://www.nstac.go.jp/services/ippan-microdata.html>) から全て入手することができる。

3.1 個々の変数の確認

一般マイクロは、質的変数と量的変数から構成され、データとともに、データレイアウトと符号表が提供されている。利用する主要な変数の状況について、分析に影響を与える可能性がある欠測や外れ値の有無なども含めて概要を把握し、データと共に提供されている文書類の記述とも矛盾がないことを確認しておくが良い。チェックすべき点としては、以下のような事項がある。

- a. データセットの大きさや変数の数
- b. 欠測の有無
- c. 質的変数の値の範囲と分布の確認
- d. 量的変数の基本統計量と分布や外れ値の確認

¹ 伊原・平澤, 2016、伊原・平澤, 2017 及び北林・伊原・平澤, 2017

ここでは特に、 c 及び d のデータの分布を確認する簡便な方法としてヒストグラムと、 d の量的変数の外れ値を検出する方法として箱ひげ図をとりあげる。

3.1.1 ヒストグラム

図1は、量的変数である消費支出 (L_Expenditure) について、そのままプロットした場合 (左) と常用対数により変換後のプロット (右) を示す。R によりヒストグラムを作成する関数は複数あるが、ここでは MASS パッケージの `truehist` 関数を使用している。

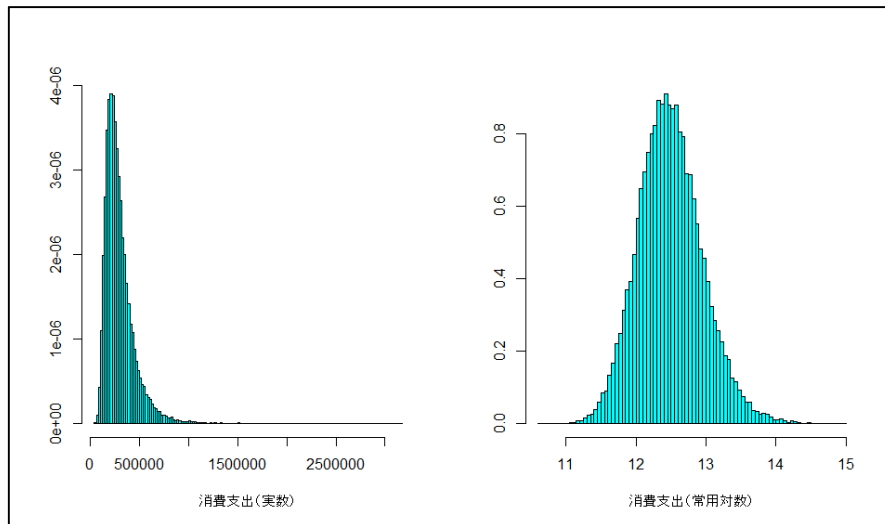


図 1. 消費支出のヒストグラム

左側のプロットでは、多くの金額データに特有な、数値の大きい右側に裾を引いた分布の形状をしていることがわかる。一方、常用対数による変換後のデータは、右側のプロット図に示すように、比較的左右対称な正規分布に近い形状となるため、元のデータ分布は対数正規分布に近いことがわかる。

3.1.2 箱ひげ図

量的変数の外れ値の確認には、箱ひげ図で表示することでデータの傾向を把握しやすい。一般に箱ひげ図は、第1四分位数から四分位範囲 (IQR: Inter Quartile Range) の1.5倍を引いた値を下限、第3四分位数に IQR の1.5倍を足した値を上限の目安として、この範囲を超える観測値を外れ値として表示する。図2.a 及び図2.b に、データの年間収入、消費支出と十大費目の12変数の対数変換の前後の値についての箱ひげ図を示す。これは、R の `base` パッケージの `boxplot` 関数による箱ひげ図で作成しており、IQR と第1・第3四分位の代わりに、ヒンジ幅とヒンジ値を使用している。

ヒストグラムと同様に、対数変換前のデータをプロットした図2.a からわかるように、これらの金額データの分布はすべて明らかに右裾が長い。これを左右対称な正規分布の形に近づけるために、常用対数で変換をした結果が図2.b である。箱ひげ図は、ひげ先を越えるものを外れ値の候補としているが、図2.b では対数変換を施してもデータの分布の裾が正規分布より明らかに長く、極端に他の大部分から離れた値が少数存在するという状況ではないことが分かる。

3.2 変数間の関係性の把握

個々の変数の状況を把握した後に、変数間の関係性に着目する。ここでは、主に量的変数に適用する散布図行列による方法と、平行座標プロットによる方法を取りあげ、最後に着目する目的変数と説明変数の間の関係性について分析を行うことのできる決定木について紹介する。演習教材作成という分析目的と作業効率を考えれば、ここで変数の組み合わせ全てをチェックする必要はなく、詳細分析につながるヒントが得られれば良い。

3.2.1 散布図行列

数量項目の概要を把握するために便利な方法の一つが、散布図行列である。例として、図 2.a 及び図 2.b と同じ一般ミクロの年間収入、消費支出と十大費目の 12 変数について図 3.a 及び図 3.b に散布図行列を示す。ここでも、常用対数変換前の値をプロットした図 3.a に加えて、常用対数で変換した値をプロットした図 3.b を作成した。対数化によりどの変数もおおむね分布が楕円に近くなることが確認できる。

散布図行列では、変数同士の相関構造の有無も確認することができる。二変数間に相関がなければ、データ分布の形は円に近いが、相関があれば楕円形に近くなり、相関が高いほど楕円の長軸と短軸の差異が大きくなっていく。ここで、散布図行列の作成は、base パッケージの `pairs` 関数を使用しており、データサイズがかなり大きいいため、描画の際には `cex` パラメータを用いてプロット点が小さくなるよう指定している。

3.2.2 平行座標プロット

この他に、質的変数にも適用可能で、相関構造を目視で簡単に確認できる方法の一つとして、平行座標プロットがある。平行座標プロットは、x 軸に対象とする変数を並べ、各変数の値を y 軸で表すように、レコード毎の値を直線で結ぶもので、多変量データの可視化に役立つ。作成時には、関係のありそうな変数同士を隣に並べるのがコツである。ただし、平行座標プロットは、散布図よりもさらに適用できるデータ量が限られる。図 4 の例は、世帯主（就業者）が 30 歳未満の世帯（1,068 レコード）を用いて、MASS パッケージ収録の `parcoord` 関数により描画しているが、変数間の関係性を特定するには明らかにレコード数が多すぎるので、利用時にはさらにデータを絞る必要があるといえる。例えば、図 4 の左側「食料」「住居」及び「光熱・水道」の 3 分類に着目すると、プロット図上部については「住居」の数値が比較的低く「食料」及び「光熱・水道」が高いグループと、「住居」の数値が高く「食料」や「光熱・水道」が低いグループの、異なる傾向の 2 グループが存在していることが確認できる。一方、データが集中しているプロット図下部については、目視では変数間の関係性を確認しづらく、さらにデータを絞る必要性を示している。

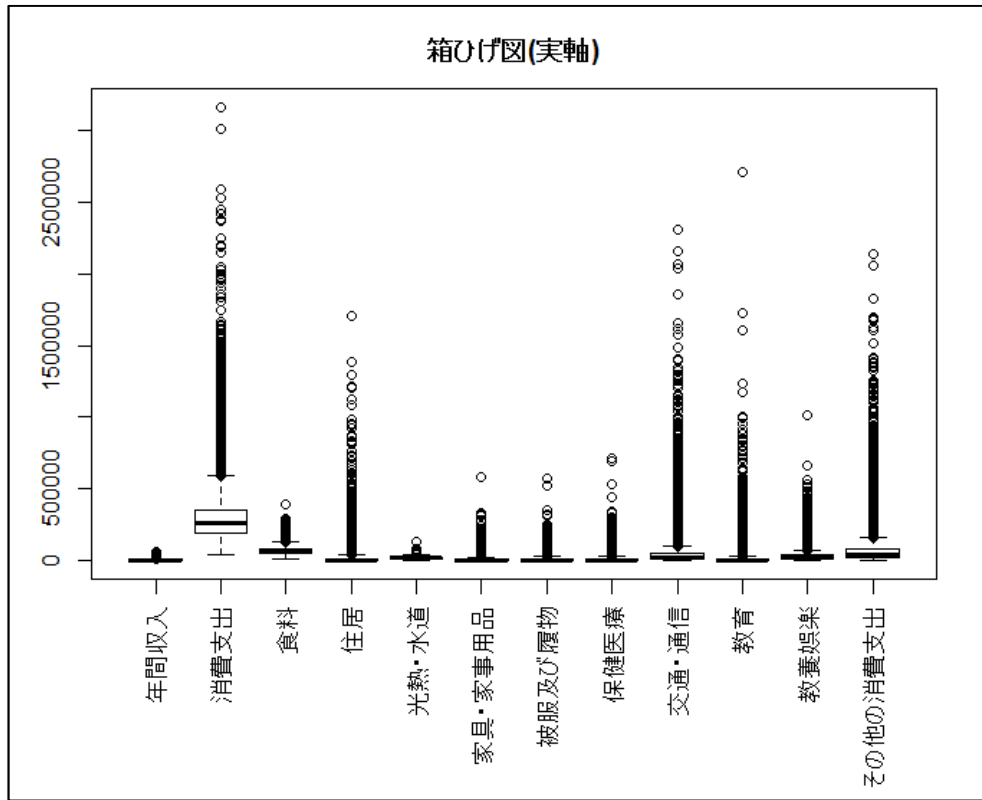


図 2.a 年間収入、消費支出及び十大費目別消費支出 (対数変換前)

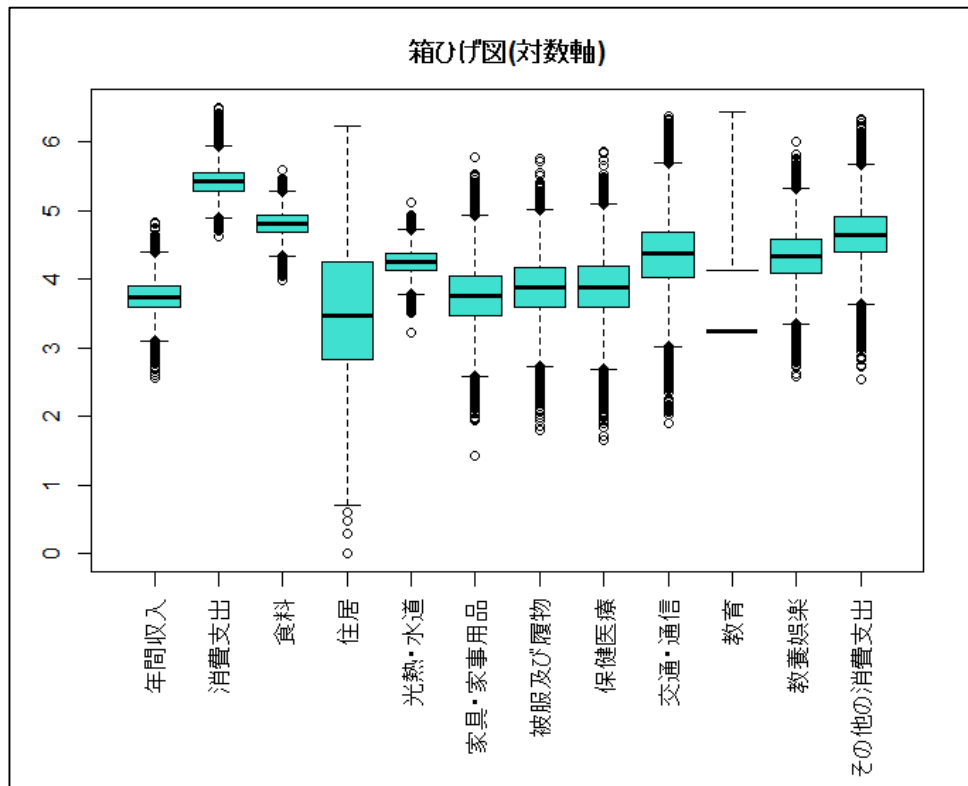


図 2.b 年間収入、消費支出及び十大費目別消費支出 (対数変換後)

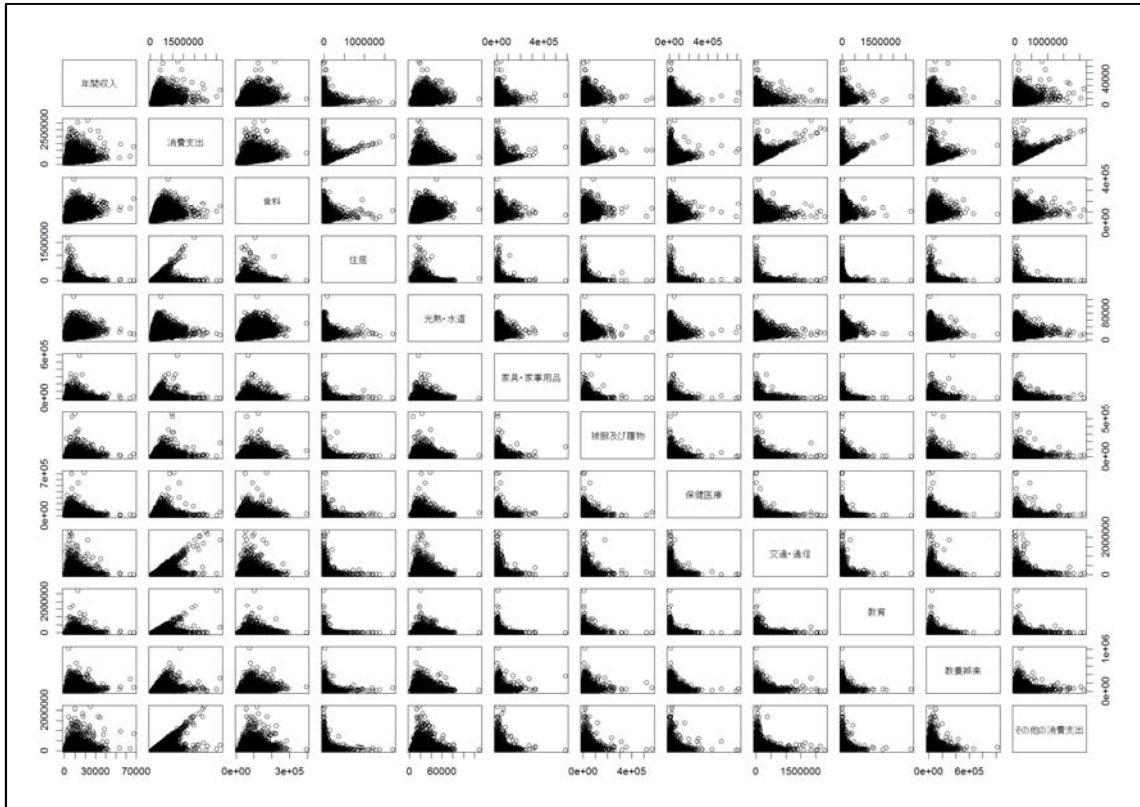


図 3.a 散布図行列 (対数変換前)

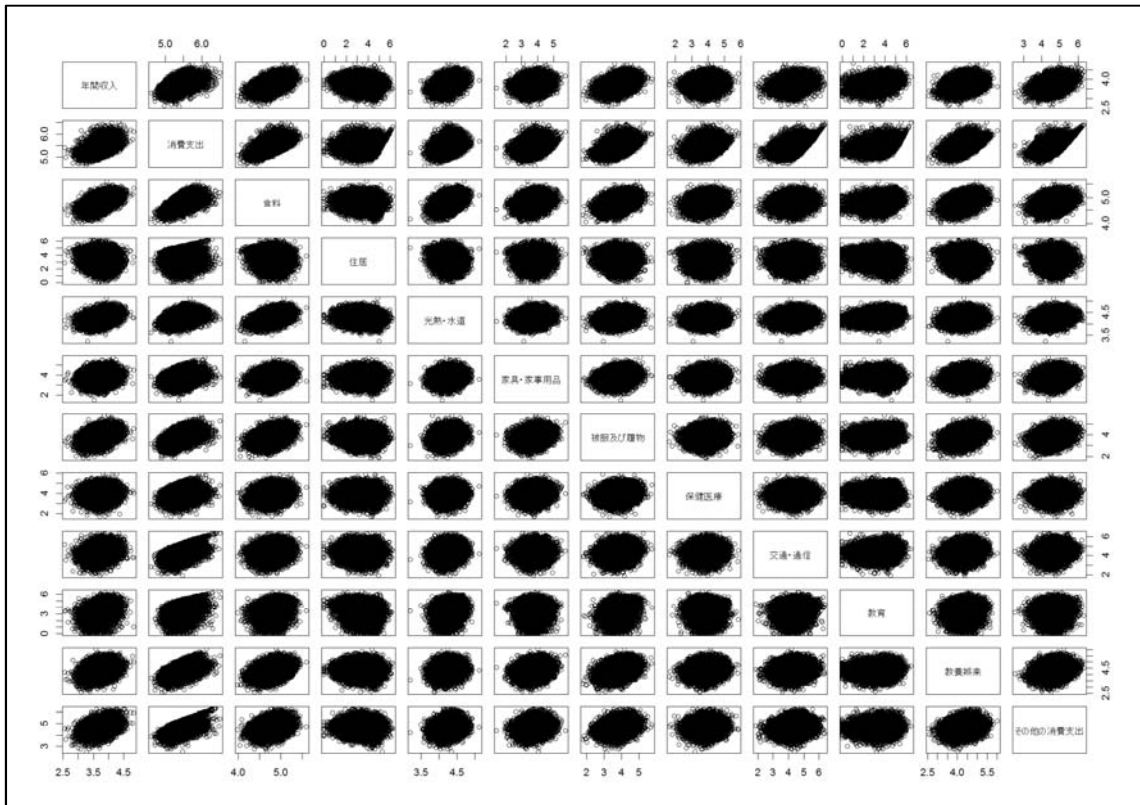


図 3.b 散布図行列 (対数変換後)

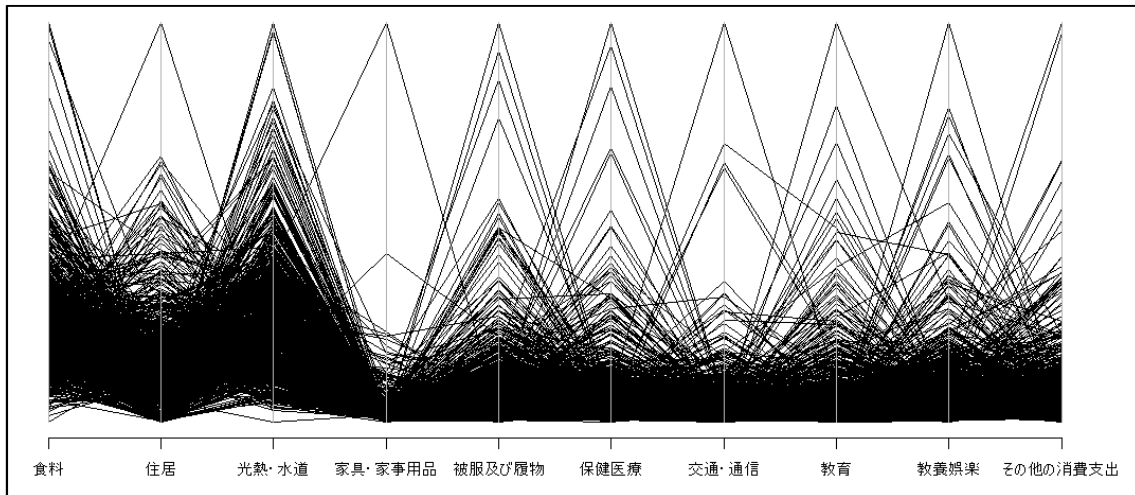


図 4. 平行座標プロット

3.2.3 決定木

決定木による分析は、データセットの中で、重要と思われる変数を目的変数として、説明変数の中から目的変数と関係のある変数を、関係が深い順番に特定することができる。決定木とは、データマイニングなどに広く使われている分析手法で、回帰木や分類木など多くの種類が存在するが、ここでは最も基本的な CART (Classification And Regression Trees) を使用する。CART は、説明変数を用いて目的変数を、分割後のデータ間の差異を最大になるように二分分割していき、さらにそれをわかりやすく可視化する機能を持つ。目的変数・説明変数共に、質的変数と量的変数の両方を取り扱うことが可能で、特に目的変数がカテゴリーの場合に分類木、連続値の場合を回帰木と呼ぶ。CART は、回帰分析などの変数選択にも利用できるが、データ分布について制約がないノンパラメトリックな方法であるという利点の一方で、データのごく一部分を置き換えるだけで構造が大きく変わる可能性があり、結果の安定性には留意が必要である。

ここでは、主要変数から、興味の対象となりそうな変数として、①年間収入と②住宅の所有を取り上げる。これらをそれぞれ目的変数として、世帯属性 7 変数と 12 の収支項目を説明変数に、`rpart` パッケージの `rpart` 関数により CART 分析を行い、`partykit` パッケージの `as.party` 関数により可視化した結果を紹介する。なお、分析にあたり、別々の変数として存在している就業者 (世帯主) の 5 歳階級別年齢 (`T_Age_5s`) と、非就業者の 65 歳未満・以上 (`T_Age_65`) という質的変数を統合し、表 1 に示す新たな年齢階級の質的変数

(`T_Age`) を新たに作成し、これら二つの質的変数の代わりに使用している。

表 1. 世帯主の年齢階級

年齢階級 T_Age		
1	(就業者)	30 歳未満
2	(就業者)	30～34 歳
3	(就業者)	35～39 歳
4	(就業者)	40～44 歳
5	(就業者)	45～49 歳
6	(就業者)	50～54 歳
7	(就業者)	55～59 歳
8	(就業者)	60～64 歳
9	(就業者)	65 歳以上
91	(非就業者)	65 歳未満
92	(非就業者)	65 歳以上

① 年間収入

年間収入による世帯の消費傾向の違いをみる。年間収入は量的変数であるため、回帰木による分析を行う。結果は図5のとおり、消費支出が306,745円で分割するとデータ間の差異を最大にすることが分かる。その次に世帯主の年齢階級、食料費の変数が選択されるという結果が得られた。消費支出が306,745円未満で、世帯主の年齢階級が30歳未満及び非就業者の世帯においては年間収入が低く、消費支出が306,745円以上で世帯主の年齢階級が40歳以上の世帯で、食料費が88,513円以上の世帯においては、年間収入が高い傾向にある。上述のとおり、世帯主年齢階級は就業・非就業の別の情報も持っており、分岐条件をみると、若年及び非就業世帯の年間収入が低い傾向があるという結果が得られた。

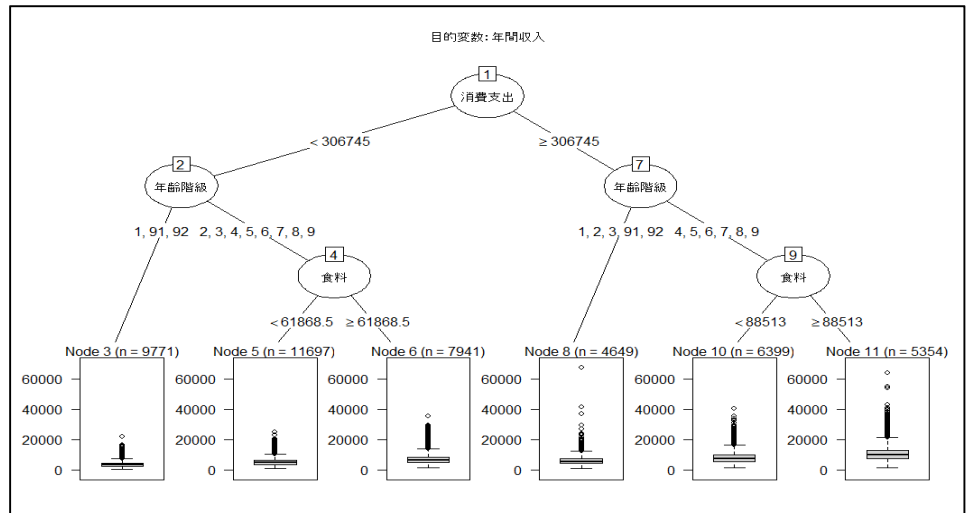


図5. 目的変数を年間収入とした CART 分析結果

② 住宅の所有

次に、「持ち家」か「借家・借間」を最も良く分離することができる変数を探す。住居の所有を目的変数とし、それ以外の変数を説明変数として分類木による分析を行った結果を、図6に示す。分割後のデータ間の差異を最大にする変数が住居費で、住居費が16,500円未満であれば、ほとんどの世帯が「持ち家」である。さらに、住居費が16,500円以上の世帯については、次に年齢階級及び就業の別の変数が選択され、(就業者)54歳以下の世帯については、「借家・借間」が全体の8割強を占める。一方、55歳以上の世帯については、消費支出による分岐があり、約266,948円を境にデータが分割されている。

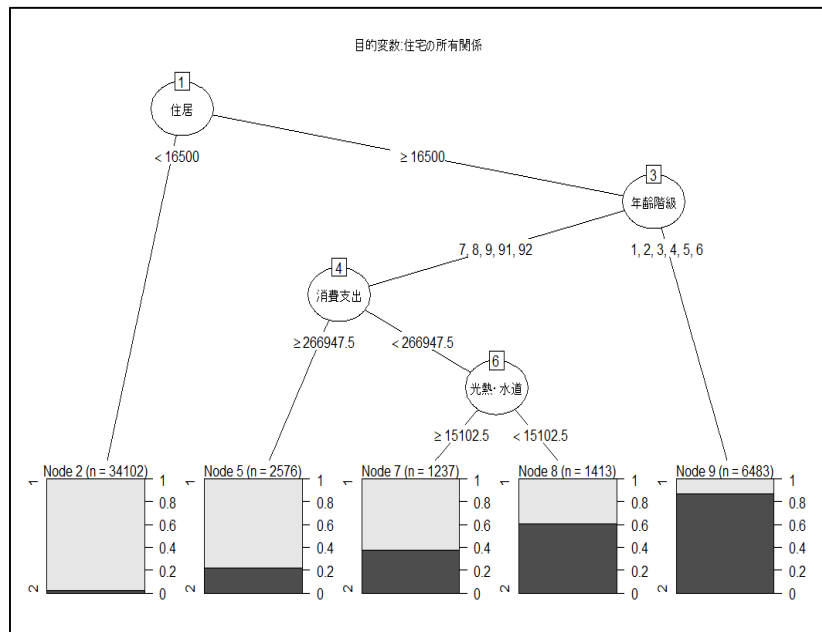


図6. 目的変数を住宅の所有関係とした CART 分析結果

3.3 詳細分析

前節までの分析において、年間収入や、住宅所有、就業・非就業と年齢階級等の間に関係が存在するという感触を元に、世帯の主食及び主菜という二つの分析テーマを設定し、より詳細な分析を試みる。

3.3.1 主食についての分析

世帯の収支項目に含まれる穀類は、「米」、「パン」、「めん類」、「他の穀類」の内訳を持っている。ここでは、内訳の中で最も消費額の高いものを、世帯の主食とみなして、主食の異なる世帯の違いについての分析例として、エンゲル係数や世帯主の年齢階級に着目した結果を紹介する。

まず、主食別の世帯数、世帯比率と穀類の合計金額、平均金額及び金額比率について集計した結果を表2に示す。なお、合計金額(千円)は、「全世帯の合計金額÷1000」、平均金額(円)は、「全世帯の合計金額÷全世帯数」で算出した。全体の80%近い世帯が、米の消費額が最も高く、パンが主食の世帯は、全体の20%弱という結果が得られた。

表2. 主食別世帯数、主食別世帯比率、合計金額、平均金額及び金額比率

	米	パン	めん類	他の穀類
世帯数	36400	8615	795	1
世帯比率(%)	79	19	2	0
合計金額(千円)	154998	99435	58927	15147
平均金額(円)	3383	2171	1286	331
金額比率(%)	47	30	18	5

さらに、世帯の主食別にエンゲル係数を比較する。エンゲル係数とは、支出総額に占める食料支出額の割合であり、世帯の所得水準が高くなるにつれ、エンゲル係数が減少するという法則はよく知られている。この他、エンゲル係数の変化の要因として、食品価格の上昇や世帯の就労人員が増えたことによる費用の変化、食費以外の買い控えなどが挙げられる。エンゲル係数は、「食料÷消費支出×100」として算出した。食料には、「穀類」、「魚介類」、「肉類」、「乳卵類」、「野菜・海藻」、「果物」、「油脂・調味料」、「菓子類」、「調理食品」、「飲料」、「酒類」及び「外食」が含まれる。主食別のエンゲル係数の中央値と平均値に着目した棒グラフを図7に示す。主食が米の世帯のエンゲル係数は、他の主食の世帯よりも大きい傾向があるという結果が得られた。

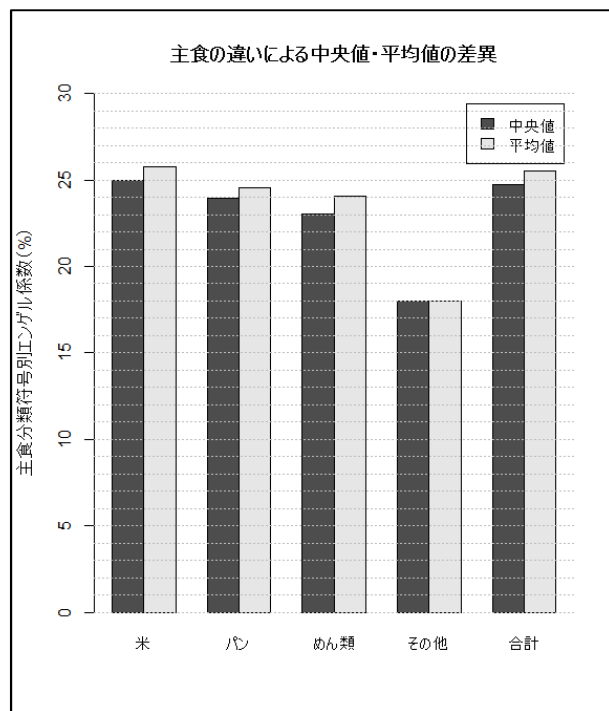


図7. 主食の違いによる中央値・平均値の差異

3.3.2 主菜についての分析

主食の分析と同様に、魚介類と肉類の消費額の多い方を世帯の主菜とする。まず、魚介類と肉類、それぞれの消費金額と世帯主の年齢との関係性について分析し、就業者（世帯主）の年齢階級別の魚と肉の平均消費支出を表3に示す。結果から、49歳以下の世帯では肉の消費金額が多く、50歳以上の世帯では魚の消費金額が多くなっている。45歳～49歳と50～54歳の年齢階級を境に魚と肉の世帯における消費金額が逆転していることが折れ線グラフにより可視化した図8により読み取ることができる。ちなみに、世帯主が就業者でない世帯では、魚と肉の平均消費額はそれぞれ6,900円と4,800円である。

表 3. 年齢階級別主菜の平均消費金額

世帯主 年齢階級	30歳未満	30～34歳	35～39歳	40～44歳	45～49歳	50～54歳	55～59歳	60～64歳	65歳以上
魚	3149	4081	4642	6109	6539	7413	7163	8185	8275
肉	3959	4934	5711	6573	7005	7053	6592	6008	5382

(円)

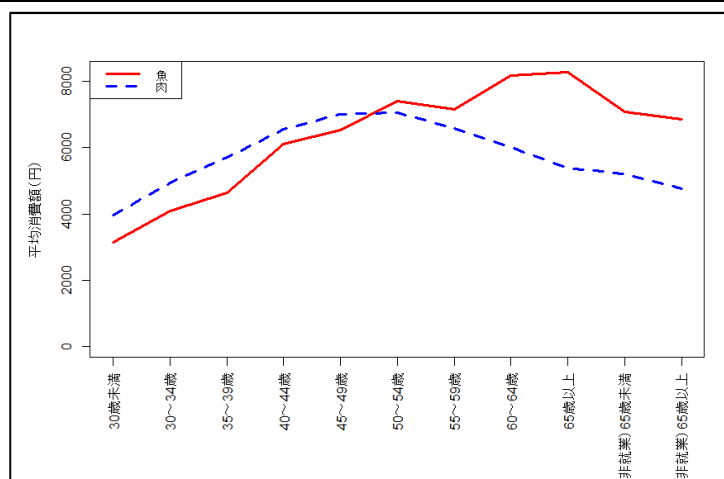


図 8. 年齢階級別主菜の平均消費金額

4. 一般マイクロの詳細品目版と擬似マイクロとの相違点

一般マイクロの詳細品目版と擬似マイクロには、いくつかの相違があり、収録変数の相違と、作成方法に起因する相違に分けることができる。本節では、これらの相違点とその影響について整理する。

4.1 収録変数の相違

二つのマイクロデータについては第2節で詳しく述べたが、収録変数の数自体は、一般マイクロの方が多。ただし、分析時により重要性が高いのは、世帯属性情報である。これは、擬似マイクロの14変数に対し、一般マイクロが7変数で、さらに後者はカテゴリの数が就業者の年齢階級以外には2つしかない。比較のために、すでに公開が終了している擬似マイクロの世帯属性の符号表を付録3、量的変数である用途分類項目の一覧を付録4として本稿の末尾に添付した。

一般マイクロについては、非就業者の年齢階級が65歳未満と以上の2区分、世帯人員数が2人と3人以上の2区分であり、3.3.1節に主食についての詳細分析を例示したが、これらの変数についてより詳細な区分を持つ擬似マイクロでは、主食の違う世帯の特徴をさらに掘り下げることも可能であった。図9の結果から、擬似マイクロでは世帯人員が4人までであれば、人員が増えるほ

ど主食がパンの世帯が増えるが、それ以上は人員が増えるほど米の世帯の比率が増加するという傾向がみられる。また、図10により、擬似マイクロでは世帯主の年齢が45歳未満と以上で、主食パンの世帯数と米の世帯数が逆転していくことがわかる。

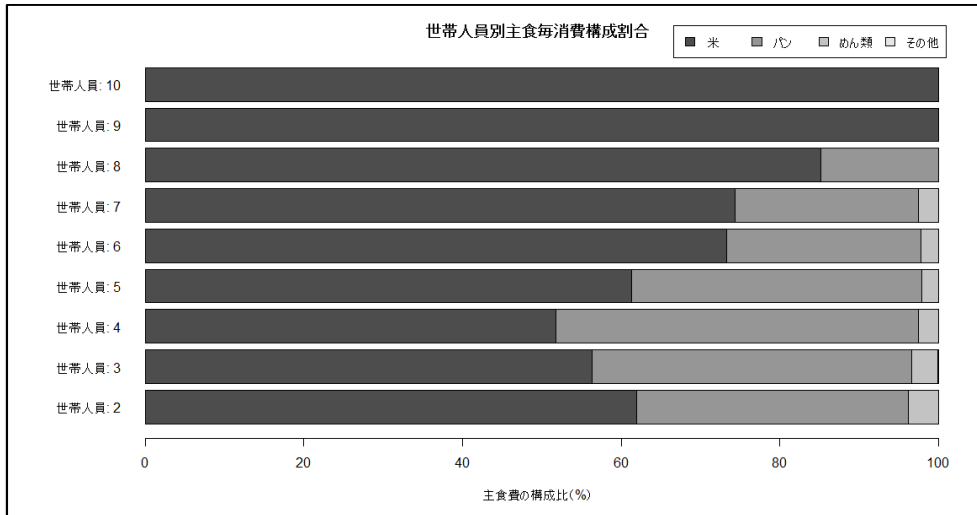


図9. 世帯人員別主食の消費金額の構成比 (擬似マイクロ)

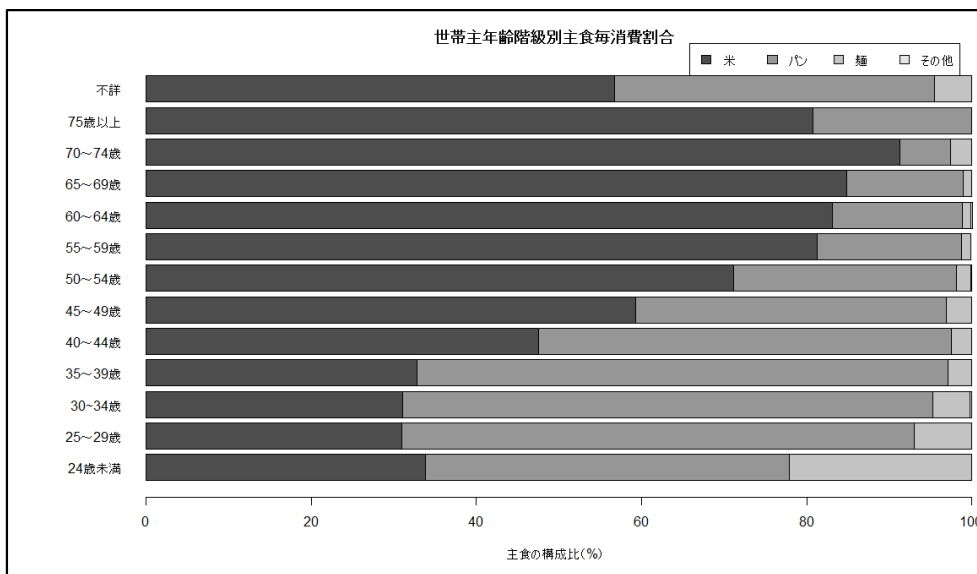


図10. 世帯主年齢階級別主食の消費金額の構成比 (擬似マイクロ)

4.2 作成方法に基づく相違

一般マイクロ、擬似マイクロともに、調査のマイクロデータの秘匿性を確保するために、一旦集計表を介して乱数を発生させて作成している。ただし、擬似マイクロは世帯単位で変数間の相関構造を反映し作成されているのに対し、一般マイクロは十大費目について世帯属性別に相関構造が反映されているが、その内訳となる詳細品目は、世帯属性が同じであれば統計的に差異がない。このため、世帯属性を絡めずに品目同士の消費額の関係性について分析する場合、必ずしも意味のある結果が出るとは限らない。

また、一般マイクロの作成にあたり、世帯属性別のゼロ値情報が利用可能でないため、特定の世帯だけが消費しそれ以外の世帯は通常購入しないような品目について、ゼロ値か必ずしも実態に即して反映されていないことに留意が必要である。具体例を挙げると、一般マイクロのアルコールの種類には「清酒」、「焼酎」、「ビール」、「ウイスキー」、「ワイン」、「発泡酒・ビール風アルコール飲料」及び「他の酒」の7種類がある。これらの酒別平均消費金額は、表4のとおり、最もビールの消費金額が多くウイスキーの消費金額は少ない。しかし、いずれの酒別についてもゼロ値の世帯がなく、全世帯が全てのアルコールを消費しており、実態との乖離がある点に留意が必要である。

表4 一般マイクロのアルコール種別消費金額の基本統計量

	清酒	焼酎	ビール	ウイスキー	ワイン	発泡酒・ビール風アルコール飲料	他の酒
最小値	21	13	38	1	7	25	5
第1四分位	350	305	810	49	110	252	84
中央値	512	440	1158	80	168	362	125
平均値	598.4	501.7	1324	94.8	201.2	411.8	145.5
第3四分位	748	626	1644	122	250	514	183
最大値	6754	6033	10326	1111	3916	3258	1152

(円)

分析を進め、世帯の購入で最も消費金額が多いアルコールに着目すると、約92%の世帯が主にビールを購入し、次いで清酒(約5%)、焼酎(約2%)の順になる。

さらに、アルコールの消費金額について、酒別の消費金額に影響する変数を探するため、酒別を目的変数、その他の主要20変数を説明変数としてCARTにより分析を行った結果、年齢階級という結果が得られた。さらに非就業者の世帯では3大都市圏か否かが影響するが、ほとんどの世帯においてビールの消費額が多いため、目的変数の分離は二番目に消費額の多い清酒までが限界であり、これはゼロ値が再現されていない点が分析に悪影響を及ぼしている可能性がある。

5. まとめ

社会人や大学生の基礎教育の目的で、統計分析の演習などに自由に利用することのできるマイクロデータは、ある程度以上の規模のものが少なく、特に公的統計に基づくデータはユーザーのニーズが高い一方で、第2節に述べたように利用可能なものが限られている。教育分野での統計教育及びデータリテラシーに関するニーズについては渡辺(2013)に詳しいが、公的部門においても、平成28年12月の経済財政諮問会議で決定された「統計改革の基本方針」において、エピソード・ベースの政策立案から、エビデンス・ベースの政策立案への転換が謳われ、「エビデンス/証拠」となる統計等の整備・改善や、ユーザーニーズを反映した統計の重要性が確認された。さらに、平成29年5月の統計改革推進会議における最終とりまとめでは、国・地方の職員一般に求められるデータ・リテラシーとして、データに基づく合理的な思考により課題を解決する能力が明記されている。

データ・リテラシーの向上には、良い学習教材が不可欠であり、良い教材を作成する上で、良い分析データは非常に重要な役割を果たす。近年の、擬似マイクロ及び一般マイクロの作成・公表は、マイクロデータの分析能力向上に資することのできる大きな流れであり、今後、一般に広く利

用できる匿名データあるいはパブリックユースファイルを作成するという動き（奥積, 2017）とともに、より一層ユーザーニーズに沿うマイクロデータの作成が期待される。

統計分析の教材を作成する目的を持つユーザーとして、望ましいマイクロデータ分析用データの要件を挙げれば、公表されている統計表との一致性よりも、実際の世帯に存在しうる変数の値の組み合わせであることと、受講生の興味を引くようある程度深掘りできる分析テーマが設定できることが重要である。実際の世帯に存在しうる変数の値の組み合わせという点で、例えばある調査項目の数値が大きいレコードは、別のある調査項目の値が小さい傾向があるといったような変数間の相関の情報は重要であり、問題の原因を追究できる分析テーマのためには調査客体の属性情報の豊富さが大いに関係する。

一般マイクロは利用登録をなくして自由にダウンロード・利用が可能になったという点で利便性がさらに向上している。一般マイクロは、その作成方法上の制約により、留意が必要な点はあるが、総務省統計研究研修所の平成29年度特別コース「マイクロデータ分析の基礎」において、演習教材の一部（和田・河野, 2017）に一般マイクロを使用し、持ち帰りの復習等、再利用も可能な点を含め、研修生からは良い評価が得られている。

将来的には、調査客体が世帯のみでなく企業や事業所にも拡充され、時系列やパネルデータの形でのデータセットが作成・公表されることを期待したい。

参考文献

- [1] 伊原一, 平澤 鋼一郎 (2016) 一般用マイクロデータの試作データ (詳細版) 作成手法～平成21年全国消費実態調査に基づく擬似データ～, 「匿名データ等利用推進ワークショップ」発表資料, 一橋大学経済研究所附属社会科学統計情報研究センター・神戸大学経済経営研究所共催, 平成28年12月20日, 一橋大学 国立西キャンパス 経済研究所.
http://www.nstac.go.jp/services/society_paper/28_06_01.pdf
- [2] 伊原一, 平澤 鋼一郎 (2017) 一般用マイクロデータ詳細品目版の作成手法 (仮) ～平成21年全国消費実態調査に基づく擬似データ～, 2017年度統計関連学会連合大会講演報告集, pp. 125.
- [3] 奥積雅彦 (2017) ミクロデータの提供の新たな展開方向の一考察, 製表技術参考資料 35, 独立行政法人統計センター.
- [4] 北林三就, 伊原一, 平澤 鋼一郎 (2017) 一般用マイクロデータ詳細品目版の概要 —平成21年全国消費実態調査に基づく擬似データ—, 経済統計学会第61回全国研究大会報告要旨集, pp.35-36.
- [5] 山口幸三, 伊藤伸介, 秋山裕美 (2013) 教育用擬似マイクロデータの作成—平成16年全国消費実態調査を例として—, 統計学, 2013, No. 104, pp.1-15.
- [6] 渡辺美智子 (2013) 知識基盤社会における統計教育の新しい枠組み～科学的探究・問題解決・意思決定に至る統計思考力～, 日本統計学会誌, 第42巻, 第2号, pp. 253-271.
- [7] 和田かず美, 河野真理子 (2017) 一般用マイクロデータ詳細品目版によるデータ分析編, 総務省統計研究研修所平成29年度特別コース「マイクロデータ分析の基礎」テキスト第II分冊.
<https://www.slideshare.net/k.wada/2017-83192029>

付録1 本文中で作成した図のRコード

```
setwd("D:/K01/imicro")          #ディレクトリの指定

#一般用マイクロデータ詳細品目版の読み込み
dat <- read.table("ippan_2009zensho_s_dataset.csv", header=TRUE, sep=";", skip=8,
  colClasses=c(rep("factor",7), rep("numeric",423)))
attach(dat)

items <- read.csv("ippan_items.csv", header=TRUE)
  # 符号表から作成する変数ラベル (日本語名)。詳細は演習教材を参照

#世帯主の年齢階級 T_Age 作成
dat$T_Age <- 0 # データフレーム dat に T_Age という変数を作成
dat$T_Age[which(T_Syuhi==1)] <- as.character(T_Age_5s[which(T_Syuhi==1)])
dat$T_Age[which(T_Syuhi==2 & dat$T_Age_65==1)] <- "91"
dat$T_Age[which(T_Syuhi==2 & dat$T_Age_65==2)] <- "92"
dat$T_Age <- as.factor(dat$T_Age) #型変換
detach(dat) # 新変数 T_Age もデータフレーム名 dat が省略で
attach(dat)

#=====
#消費支出のヒストグラム
library(MASS)          #MASSパッケージを利用
par(mfcol=c(1,2))
truehist(L_Expenditure,xlab="消費支出 (実数) ")
truehist(log(L_Expenditure),xlab="消費支出 (常用対数) ")

dt1 <- dat[,9:20] #年間収入、消費支出および十大費目のデータ作成
colnames(dt1) <- items$Japanese[9:20]
head(dt1)

#=====
#年間収入、消費支出及び十大費目の箱ひげ図
par(mar=c(8,3,3,2)) # グラフィック画面の余白の設定 マージン調整
boxplot(dt1, las=3, main="箱ひげ図(実軸)") #las=3 で軸ラベルを縦に
boxplot(log10(dt1), las=3, main="箱ひげ図(対数軸)", col="turquoise")

#=====
#散布図行列
png(filename="散布図行列_実軸.png", width=2339, height=1654, pointsize = 32)
plot(dt1) ; dev.off()
png(filename="散布図行列_対数.png", width=2339, height=1654, pointsize = 32)
plot(log10(dt1)) ; dev.off()

#=====
#平行座標プロット
#世帯主の年齢階級が (就業者) 30 歳未満限定し、教育に 0 値があるため対象から除く
dt <- dat[T_Age_5s==1,11:20]
idx <- which(dt$Education==0)
colnames(dt) <- items$Japanese[11:20]
par(mar=c(3,0,0,2))
parcoord(dt[-idx,],)

#=====
#決定木
dt2 <- cbind(dat[,1:20], dat[,431]) #世帯属性及び収支項目と世帯主の年齢階級にデータを限定する
colnames(dt2) <- c(as.character(items$Japanese[1:20]),"年齢階級")
require(rpart) # rpart 関数 最初はインストールが必要
require(partykit) # 決定木の可視化に必要な 最初はインストールが必要
rm1 <- 6:8 # 第 6 変数年齢階級 2, 7 変数年齢階級 1, 第 8 変数乗率
# 住宅の所有・非所有 (目的変数)
(ct1 <- rpart(dt2$住宅の所有関係~, dat=dt2[, -rm1], method="class"))
plot(as.party(ct1), main="目的変数:住宅の所有関係")
```



```

# 年間収入 (目的変数)
ct3 <- rpart(dt2$年間収入~, dat=dt2[, -rm1], method="anova")# method を"anova" (回帰木)
plot(as.party(ct3), main="目的変数：年間収入")

#=====
#主食の違いによる分析
# 主食符号の作成 主食インデックス (g1-g4) の作成
g1 <- which(E003 >= E004 & E003 >= E005 & E003 >= E006)
g2 <- which(E004 > E003 & E004 >= E005 & E004 >= E006)
g3 <- which(E005 > E003 & E005 > E004 & E005 >= E006)
g4 <- which(E006 > E003 & E006 > E004 & E006 > E005)
## 主食符号 (f.MD) セット
f.MD <- rep(NA, dim(dat)[1])
f.MD[g1] <- 1; f.MD[g2] <- 2; f.MD[g3] <- 3; f.MD[g4] <- 4
table(f.MD, useNA="always") # 主食符号別世帯数

# 主食符号別の金額比率を計算してみる
t1 <- matrix(NA, nr=5, nc=4) # 5行4列
colnames(t1) <- items$Japanese[23:26] # 漢字の項目名を付与
rownames(t1) <- c("世帯数", "世帯比率%", "合計金額(千円)", "平均金額(円)", "金額比率%")
t1[1,] <- table(f.MD) # 世帯数
t1[2,] <- t1[1,] / sum(t1[1,]) * 100 # 世帯比率
t1[3,] <- apply(dat[,23:26], 2, sum) / 1000 # 合計金額[千円]
t1[4,] <- apply(dat[,23:26], 2, mean) # 平均金額[円]
t1[5,] <- t1[3,] / sum(t1[3,]) * 100 # 比率計算
round(t1)

#主食別に算出して縦ベクトルを横に結合する
out1 <- cbind(summary(Food[g1]/L_Expenditure[g1]),summary(Food[g2]/L_Expenditure[g2]),
summary(Food[g3]/L_Expenditure[g3]),summary(Food[g4]/L_Expenditure[g4]),
summary(Food / L_Expenditure)) * 100
colnames(out1) <- c("米", "パン", "めん類", "その他", "合計")
round(out1, digits=3) # 小数点第3位で四捨五入
record <- c(t1[1,], sum(t1[1,])) # 主食別のレコード数及び合計
out1 <- rbind(out1, record) # 縦に結合

#=====
# 棒プロット(主食分類符号別エンゲル係数の中央値と平均値)
barplot(out1[3:4,], beside=TRUE, legend=c("中央値", "平均値"), ylim=c(0,30),
ylab="主食分類符号別エンゲル係数 (%)", main="主食の違いによる中央値・平均値の差異")
abline(h=0:30, col="gray", lty=3)

#=====
# 主菜 (食料・魚介類・肉類) の分析
# 魚と肉の消費、年齢階級別
fishmeat.age <- matrix(NA, nr=2, nc=11) # 作成する表のイメージ
colnames(fishmeat.age) <- c("30歳未満", "30~34歳", "35~39歳", "40~44歳", "45~49歳",
"50~54歳", "55~59歳", "60~64歳", "65歳以上", "(非就業) 65歳未満", "(非就業) 65歳以上")
rownames(fishmeat.age) <- c("魚", "肉")
fishmeat.age[1,] <- tapply(E007, T_Age, mean)
fishmeat.age[2,] <- tapply(E012, T_Age, mean)
fishmeat.age # 表示

#=====
# 魚と肉の消費額、年齢階級別の折れ線グラフ
par(mar=c(8,5,3,2))
plot(0, 0, type="n", xlim=c(1, ncol(fishmeat.age)), ylim=c(0, max(fishmeat.age)), xaxt="n",
xlab="", ylab="平均消費額 (円)") # x軸メモリをxaxtで消す
axis(1, at=1:11, labels=colnames(fishmeat.age), las=3) # x軸の各ラベル名
lines(fishmeat.age[1,], lwd=3, lty=1, col="red") # 魚の折れ線グラフ
lines(fishmeat.age[2,], lwd=3, lty=2, col="blue") # 肉の折れ線グラフ
legend("topleft", legend=rownames(fishmeat.age), lty=c(1,2), lwd=3, col=c("red", "blue"))

```

付録2 一般用マイクロデータ符号表（世帯属性及び十大費目のみ）

政府統計コード		00200564	実施時期	作成日	2017/6/9			
統計調査名		平成21年全国消費実態調査	集計区分	全世帯(二人以上の世帯)	ファイル名			
調査種別		一般用マイクロデータ(十大費目)	備考・補足事項	(照会先等)				
コード体系		Shift_JIS						
項目総数		20						
文字列引用符								
行番号	項目名	階層番号	項目番号	種別	変数名	符号	符号内容	備考
1	3大都市圏か否か	1	1	3City		1	3大都市圏	
2						0	3大都市圏以外	
3	世帯人員	1	2	T_SevJnin		2	2人	
4						3	3人以上	
5	就業人員	1	3	T_SyuJnin		1	1人以下	
6						2	2人以上	
7	住宅の所有者関係	1	4	T_JuSyoyu		1	持ち家	
8						2	借家・借間	
9	就業・非就業の別	1	5	T_Svwhi		1	就業	
10						2	非就業	
11	年齢階級2	1	6	T_Age_5s		1	(就業者)30歳未満	
12						2	(就業者)30～34歳	
13						3	(就業者)35～39歳	
14						4	(就業者)40～44歳	
15						5	(就業者)45～49歳	
16						6	(就業者)50～54歳	
17						7	(就業者)55～59歳	
18						8	(就業者)60～64歳	
19						9	(就業者)65歳以上	
20						0	就業者以外	
21	年齢階級1	1	7	T_Age_65		1	65歳未満	
22						2	65歳以上	
23	集計用乗車	1	8	Weight		0～	集計用乗車	
24	年間収入	1	9	Y_Income		0～	単位:千円	
25	消費支出	1	10	L_Expenditure		0～	単位:円	
26	食料	1	11	Food		0～	単位:円	
27	住居	1	12	Housing		0～	単位:円	
28	光熱・水道	1	13	LFW		0～	単位:円	
29	家具・家事用品	1	14	Furniture		0～	単位:円	
30	娯楽及び服飾	1	15	Clothes		0～	単位:円	
31	保健医療	1	16	Health		0～	単位:円	
32	交通・通信	1	17	Transport		0～	単位:円	
33	教育	1	18	Education		0～	単位:円	
34	娯楽娯楽	1	19	Recreation		0～	単位:円	
35	その他の消費支出	1	20	OL_Expenditure		0～	単位:円	

付録3. 教育用擬似マイクロデータ符号一覧 (世帯属性のみ)

項目名	変数名	符号	符号内容		
世帯区分	SetaiKubun	1	勤労		
		2	勤労以外		
		3	無職		
世帯票					
世帯人員	SetaiJinin	△2～	2人～		
有業人員	ShuugyouJinin	△1～	1人～		
		VV	不詳		
現住居等に関する事項					
住居の構造	Kouzou	1	木造		
		2	防火木造		
		3	鉄骨・鉄筋コンクリート造		
		4	その他(ブロック造り、レンガ造りなど)		
		V	不詳		
住居の建て方	Tatekata	1	一戸建		
		2	長屋建		
		3	共同住宅(1・2階建)		
		4	共同住宅(3～5階建)		
		5	共同住宅(6～10階建)		
		6	共同住宅(11階建以上)		
		7	その他		
		V	不詳		
住居の所有関係	Shoyuu	1	持ち家(世帯員名義)		
		2	持ち家(その他名義)		
		3	民営賃貸住宅(設備専用)		
		4	民営賃貸住宅(設備共用)		
		5	県市区町村営賃貸住宅		
		6	都市再生機構・公社等賃貸住宅		
		7	社宅・公務員住宅(借上げ含む)		
		8	借間		
		V	寮・寄宿舎		
世帯員に関する事項					
世帯主					
性別	S1_Sex	1	男		
		2	女		
年齢5歳階級	S1_Age	△1	24歳未満		
		△2	25～29歳		
		△3	30～34歳		
		△4	35～39歳		
		△5	40～44歳		
		△6	45～49歳		
		△7	50～54歳		
		△8	55～59歳		
		△9	60～64歳		
		△10	65～69歳		
				11	70～74歳
				12	75歳以上
		VV	不詳		
就業・非就業の別	S1_Shugyou	1	就業		
		2	うちパート		
		3	非就業		
		4	うち仕事を探している		
		V	不詳		

項目名	変数名	符号	符号内容
企業区分	S1_KigyokuKubun	1	民営
		2	自営
		3	官公
		△	非就業
		V	不詳
企業規模	S1_KigyokuKibo	1	1～4人
		2	5～29人
		3	30～499人
		4	500～999人
		5	1000人以上
		△	非就業又は官公
		V	不詳
産業符号	S1_Sangyou	△1	農業
		△2	林業
		△3	漁業
		△4	鉱業
		△5	建設業
		△6	製造業
		△7	電気・ガス・熱供給・水道業
		△8	情報通信業
		△9	運輸業
		10	卸売・小売業
		11	金融・保険業
		12	不動産業
		13	飲食店・宿泊業
		14	医療・福祉
		15	教育・学習支援業
		16	複合サービス事業
		17	サービス業(他に分類されないもの)
		18	公務(他に分類されないもの)
		19	その他(非就業を含む)
		VV	不詳
職業符号	S1_Shokugyou	△1	常用労務作業
		△2	臨時及び日々雇労務作業
		△3	民間職員
		△4	官公職員1
		△5	官公職員2
		△6	商人及び職人
		△7	個人経営者
		△8	農林漁業従事者
		△9	法人経営者
		10	自由業者
		11	その他
		12	無職
集計用乗率	Weight	△△△△△ △△△ 0.00001～	△△△△△△△△0.00001～

付録4. 教育用擬似マイクロデータ量的変数一覧（用途分類）

項目名	階層	変数名	単位	項目名	階層	変数名	単位	項目名	階層	変数名	単位
年間収入	2	Youto001	千円	生鮮果物	7	Youto062	円	交通	6	Youto123	円
収入総額	2	Youto002	円	果物加工品	7	Youto063	円	自動車等関係費	6	Youto124	円
実収入	3	Youto003	円	油脂・調味料	6	Youto064	円	自動車等購入	7	Youto125	円
経常収入	4	Youto004	円	油脂	7	Youto065	円	自転車購入	7	Youto126	円
勤め先収入	5	Youto005	円	調味料	7	Youto066	円	自動車等維持	7	Youto127	円
事業・内職収入	5	Youto006	円	菓子類	6	Youto067	円	通信	6	Youto128	円
農林漁業収入	6	Youto007	円	調理食品	6	Youto068	円	教育	5	Youto129	円
家賃収入	6	Youto008	円	主食的調理食品	7	Youto069	円	授業料等	6	Youto130	円
他の事業収入	6	Youto009	円	他の調理食品	7	Youto070	円	教科書・学習参考教材	6	Youto131	円
内職収入	6	Youto010	円	飲料	6	Youto071	円	補習教育	6	Youto132	円
本業以外の勤め先・事業・内職収入	5	Youto011	円	茶類	7	Youto072	円	教養娯楽	5	Youto133	円
他の経常収入	5	Youto012	円	コーヒー・ココア	7	Youto073	円	教養娯楽用耐久財	6	Youto134	円
財産収入	6	Youto013	円	他の飲料	7	Youto074	円	教養娯楽用品	6	Youto135	円
社会保障給付	6	Youto014	円	酒類	6	Youto075	円	書籍・他の印刷物	6	Youto136	円
公的年金給付	7	Youto015	円	外食	6	Youto076	円	教養娯楽サービス	6	Youto137	円
他の社会保障給付	7	Youto016	円	一般外食	7	Youto077	円	宿泊料	7	Youto138	円
仕送り金	6	Youto017	円	学校給食	7	Youto078	円	バック旅行費	7	Youto139	円
特別収入	4	Youto018	円	住居	5	Youto079	円	月謝類	7	Youto140	円
受贈金	5	Youto019	円	家賃地代	6	Youto080	円	他の教養娯楽サービス	7	Youto141	円
その他	5	Youto020	円	設備修繕・維持	6	Youto081	円	その他の消費支出	5	Youto142	円
実収入以外の収入	3	Youto021	円	設備材料	7	Youto082	円	諸雑費	6	Youto143	円
預貯金引出	4	Youto022	円	工事その他のサービス	7	Youto083	円	理美容サービス	7	Youto144	円
保険取金	4	Youto023	円	光熱・水道	5	Youto084	円	理美容用品	7	Youto145	円
個人・企業年金保険取金	5	Youto024	円	電気代	6	Youto085	円	身の回り用品	7	Youto146	円
他の保険取金	5	Youto025	円	ガス代	6	Youto086	円	たばこ	7	Youto147	円
有価証券売却	4	Youto026	円	他の光熱	6	Youto087	円	その他の諸雑費	7	Youto148	円
株式売却	5	Youto027	円	上下水道料	6	Youto088	円	こづかい(使途不明)	6	Youto149	円
他の有価証券売却	5	Youto028	円	家具・家事用品	5	Youto089	円	交際費	6	Youto150	円
土地家屋借入金	4	Youto029	円	家庭用耐久財	6	Youto090	円	食料	7	Youto151	円
他の借入金	4	Youto030	円	家事用耐久財	7	Youto091	円	家具・家事用品	7	Youto152	円
分割払・一括払購入借入金	4	Youto031	円	冷暖房用器具	7	Youto092	円	被服及び履物	7	Youto153	円
財産売却	4	Youto032	円	一般家具	7	Youto093	円	教養娯楽	7	Youto154	円
その他	4	Youto033	円	室内装備・装飾品	6	Youto094	円	他の物品サービス	7	Youto155	円
繰入金	3	Youto034	円	寝具類	6	Youto095	円	贈与金	7	Youto156	円
支出総額	2	Youto035	円	家事雑貨	6	Youto096	円	他の交際費	7	Youto157	円
実支出	3	Youto036	円	家事用消耗品	6	Youto097	円	仕送り金	6	Youto158	円
消費支出	4	Youto037	円	家事サービス	6	Youto098	円	非消費支出	4	Youto159	円
食料	5	Youto038	円	被服及び履物	5	Youto099	円	直接税	5	Youto160	円
穀類	6	Youto039	円	和服	6	Youto100	円	勤労所得税	6	Youto161	円
米	7	Youto040	円	洋服	6	Youto101	円	個人住民税	6	Youto162	円
パン	7	Youto041	円	男子用洋服	7	Youto102	円	他の税	6	Youto163	円
めん類	7	Youto042	円	婦人用洋服	7	Youto103	円	社会保険料	5	Youto164	円
他の穀類	7	Youto043	円	子供用洋服	7	Youto104	円	公的年金保険料	6	Youto165	円
魚介類	6	Youto044	円	シャツ・セーター類	6	Youto105	円	健康保険料	6	Youto166	円
生鮮魚介	7	Youto045	円	男子用シャツ・セーター類	7	Youto106	円	介護保険料	6	Youto167	円
塩干魚介	7	Youto046	円	婦人用シャツ・セーター類	7	Youto107	円	他の社会保険料	6	Youto168	円
魚肉練製品	7	Youto047	円	子供用シャツ・セーター類	7	Youto108	円	他の非消費支出	5	Youto169	円
他の魚介加工品	7	Youto048	円	下着類	6	Youto109	円	実支出以外の支出	3	Youto170	円
肉類	6	Youto049	円	男子用下着類	7	Youto110	円	預貯金	4	Youto171	円
生鮮肉	7	Youto050	円	婦人用下着類	7	Youto111	円	保険掛金	4	Youto172	円
加工肉	7	Youto051	円	子供用下着類	7	Youto112	円	個人・企業年金保険掛金	5	Youto173	円
乳卵類	6	Youto052	円	生地・糸類	6	Youto113	円	他の保険掛金	5	Youto174	円
牛乳	7	Youto053	円	他の被服	6	Youto114	円	有価証券購入	4	Youto175	円
乳製品	7	Youto054	円	履物類	6	Youto115	円	株式購入	5	Youto176	円
卵	7	Youto055	円	被服関連サービス	6	Youto116	円	他の有価証券購入	5	Youto177	円
野菜・海藻	6	Youto056	円	保健医療	5	Youto117	円	土地家屋借金返済	4	Youto178	円
生鮮野菜	7	Youto057	円	医薬品	6	Youto118	円	他の借金返済	4	Youto179	円
乾物・海藻	7	Youto058	円	健康保持用摂取品	6	Youto119	円	分割払・一括払購入借入金返済	4	Youto180	円
大豆加工品	7	Youto059	円	保健医療用品・器具	6	Youto120	円	財産購入	4	Youto181	円
他の野菜・海藻加工品	7	Youto060	円	保健医療サービス	6	Youto121	円	その他	4	Youto182	円
果物	6	Youto061	円	交通・通信	5	Youto122	円	繰越金	3	Youto183	円