

**科学研究費助成事業 研究成果報告書**

平成 29 年 6 月 21 日現在

機関番号：62603

研究種目：挑戦的萌芽研究

研究期間：2013～2016

課題番号：25540016

研究課題名（和文）社会調査設計における新しい層化システムの開発

研究課題名（英文）Development of a new system of stratification in the sample design of social survey

研究代表者

前田 忠彦（MAEDA, Tadahiko）

統計数理研究所・データ科学研究系・准教授

研究者番号：10247257

交付決定額（研究期間全体）：（直接経費） 3,000,000円

研究成果の概要（和文）：本研究の目的は、日本の社会調査の標本設計における層化基準を再考し、標準的な全国規模の社会調査設計の際の層化方式に関する新しい指針を提案することである。従来は地方と人口規模に応じて市区町村を層別することが多かったが、本研究では調査の主要な目的変数の性格に応じ、公的統計資料など外部の変数に依拠して適切な層化変数を選定し、実際の層化を実務的観点から効率よく行うためのシステムの開発を目指した。目的変数によっては教育水準など人口規模以外の変数がより有力な層化要因となる場合もあるが、従来の層化基準は、より多くの変数に対して平均的に有効であるという意味で役に立つ変数であることが、改めて示唆された。

研究成果の概要（英文）：The purpose of the present study is to reconsider the stratification variables in a design of Japanese social survey settings, and to propose a new strategy of stratification in sample design of a nationwide social survey. Conventionally in the sample design of a nationwide survey, municipalities were stratified according to the region and population size. By contrast, this study aims to seek for appropriate stratification variables from outside resource such as official statistics, depending on the principal target variables of a survey, and intends to develop a new effective system of stratification from the practical point of view. In some cases, variables other than the population size, such as educational level of the municipality, might be effective stratification variable. However results also suggested that population size is effective for many target variables on average and should be considered useful.

研究分野：社会調査法

キーワード：サンプリング 社会調査 公的統計 層化抽出法 日本人の国民性調査 相関比 方言意識

## 1. 研究開始当初の背景

### (1) 研究の学術的背景

統計的な社会調査の二つの要素は、調査対象者の選定のプロセスと、対象者から何らかの回答を引き出す測定のプロセスである。対象者選定のプロセスの統計的スタンダードはいわゆる無作為抽出による方法である。測定のプロセスは面接、電話など調査員のインタビューに基づく方式と、回答者が自ら調査票に記入する自記式とに大別される。本研究はこのうち、従来の調査研究では見逃されがちな対象者の選定プロセス（無作為抽出に関する問題）を専ら扱う。

平成期以降の市区町村合併により、日本の地方自治体の姿は大きく変わった。自治体そのものや住民（全国レベルでは日本国民全体）を研究対象とする際の、対象自治体・住民の性格付けも変化したことを意味する。例えば 2000 年の国勢調査人口で郡部居住の成人は約 20% であるのに対し、2010 年国勢調査では 9% 程度と半減以下となっている。約 10% の人口が（行政上の区分の変更により）市部あるいは政令指定都市内の区部に分類されることになった。このことは社会調査の標本設計上は、調査項目の分布に影響が大きいことが知られている都市的性格・農村的性格などはそれぞれの自治体でおそらく不変であるにも関わらず、制度的な理由で従来とは異なる性格の層に押し込まれることを意味する。層化のために市郡や人口規模を用いる従来の方式は、この意味で効率性を失ったと考えられ、代替案が求められている。

### (2) 研究期間内の検討目標事項

このような状況に鑑みて、本研究では新しい市区町村区分に基づく、より精度の高い調査地域の層化の方式を検討する。本研究のチャレンジとして従来の市区町村レベルでの層化よりも一段階細かい町丁字のレベルで層化基準を再設定する方式を考案し、実際の社会調査における標本設計に際して、考え方の面でも依拠すべき資料の整備・実際の地点抽出アルゴリズム等の面でも使いやすい、実際のシステムを開発することを目標とした（下記 の注記も参照）。この際、調査目的の変数の性格・テーマを大まかに区分し、従来の画一的層化変数ではなく、調査テーマに応じた層化基準の設定により、統計的な精度を上げる（サンプリング誤差の少ない）標本設計につなげるための具体的な知見を得ることも目指す。

### (3) 本研究の学術的な特色

調査の二つの要素（選定と測定）のうち、測定の側面については特に調査方法論的研究の対象に取り上げられることが多いが、もう一方の選定（標本抽出）の側面については調査の精度を決定づける重要な要素であるにも関わらず、注目が集まることが少ない。本研究は平成の市区町村大合併が一段落した現下の状況において、喫緊の検討課題に正面から取り組むものといえる。

（ ）研究当初は調査設計において町丁字等の従来の方法より詳細な単位を層化の単位と考えることを目指したが、研究期間中に行った小単位の官庁統計データを利用した分析において、そうしたデータによる分析が結果の安定性に欠ける経験をいくつかしたことから、このチャレンジについては当面断念することとし、主に市区町村単位での層化に対して有効な変数を探索し、調査設計に生かすためのノウハウを検討する方向に軌道修正を行った。

## 2. 研究の目的

本研究の目的は、近年の市区町村合併の影響を考慮して、全国規模での統計的社会調査の標本設計における層化基準の再考し、標準的な社会調査設計の際の層化方式に関する新しい指針を提案することである。標準的な全国調査の設計において、従来は地方と人口規模に応じて市区町村を層別することが多かったが、市区町村合併の影響を受けて、層化効果の減殺が懸念される。本研究では調査の主要な目的変数の性格に応じ、公的統計資料など外部の変数に依拠して適切な層化変数を選定し、実際の層化を効率よく行うためのシステム（考え方・アルゴリズム面と実装面）の開発を目指すものである。調査の統計的な意味での精度を従来の画一的な基準での層化に基づく設計よりも向上させることに寄与する研究である。

## 3. 研究の方法

本研究の研究課題は大別して A. 理論的研究と B. 実務の 2 側面の検討と C. 実証研究で構成される。A の理論的検討とは、層化の有効性を評価する指標を本研究の文脈に合わせて検討することであり、B の実務的な検討とは、上記の理論的検討を踏まえて現実に使用可能な層化変数のうちどの変数が効果を持つのかという問題を、層化変数群 X と目的変数群 Y の組み合わせという観点から、X として実際の官庁統計資料等と Y として社会調査データを利用して検討することを意味する。C の実証研究とは A や B の検討を生かして実際の既存社会調査データを解析し、また本課題内で新たに Web 調査によるデータ取得を行って、層化の効果の検討を行うこと、また関連した多変量解析などを行うことを指す。

A の理論面は平成 25 年度～26 年度、B の実務面は平成 26～28（補助期間を延長）、C の実証研究は延長期間である平成 28 年度が主な研究期間であったと総括できる。

理論面での検討は基本的に従来の標本調査法の枠内で検討することができる話題であるが、目的変数が単変量である場合の理論を目的変数が多数ある多変量の場合にどのように拡張するのが合理的か、またそれを踏まえた有効な層化変数の選択方法を考案する

ことが研究方法の中心で、机上の計算が主な手段である。

実務面の検討は、官庁統計等あるいは外部の何らかの資料（市区町村単位、町丁字等の単位、あるいはその中間の単位である選挙の投票区等の単位）のデータと、実際の社会調査データを準備し、(a)そのデータの中から適当な層化変数 $X$ を選んでその変数自体の性質を分析することと、(b)層化の候補変数を複数用意し、それらの変数によって事後的に層化した際に社会調査データの中に含まれる目的変数 $Y$ に対してどの程度の層化の効果が得られるのかを検討する、という手段による。

C の実証研究としては、既に保有している（統計数理研究所が実施した）面接調査等（例えば日本人の国民性調査、国民性に関する意識動向調査）のデータ、あるいは方言学者との共同研究として別途実施した方言の使用意識を主な調査内容とする全国調査（面接やWebによる）を分析対象とした他、本課題の中で、平成 28 年度中に新たな Web 調査を実施してデータを取得した。

平成 28 年度に実施した Web 調査は、地域の層化の効果を検証するための実データを得ることを目的とし、専門の調査会社に委託して実施した。主な仕様は次の通りである：

調査対象：日本全国の 20 歳以上男女

回収数：回収実数 7384 名

対象者数の割付等：性・年齢層・地域ブロックの要因を組み合わせ設定した 144 セルに、直近の国勢調査の 20 歳以上人口に比例した目標回収数を設定し、各セルの目標数の充足を以て調査完了とする。

調査内容：約 35 項目で、地域活動へ参加、地域環境評価、郷土に対する意識、方言意識等の、地点間の異質性が高いことが期待される項目を中心とした約 35 項目と、対象者の属性等の約 10 項目程度を含む 45 項目程度。

#### 4. 研究成果

##### (1) 層化の効果に関する理論面での成果

層化の効果を評価するための方法を考察した。残念ながら基本的に従来の標本調査理論の枠を超える検討はできなかったが、本研究での立場を述べておく。層化の効果は基本的に単変量的には（目的変数に関する層間の平均値の分離の程度を表す）相関比と呼ばれる統計量で評価される。目的変数が多変量的である場合にこれをどう拡張するかが問題となるが、(1)多変量の目的変数が全て量的変数である場合には第一主成分を計算し、それに対して相関比を評価する、(2)多変量の平均値の分離度の指標であるウィルクスの

統計量などを使う、(3)個々の変数に関する相関比の幾何平均等を用いる、などを考えた。理論的には(1)と(2)は魅力的であるが、(2)は実用上の意味が分かりにくいこと、(1)は相関が強い変数の組み合わせ以外には有効と思われないこと、などから、現実

的には(3)のような平凡な指標で評価しておくのが安全であろう。

なお社会調査の目的変数はしばしばカテゴリカルな変数であり、この場合の層平均の分離の程度は更に評価が難しい。1 個のカテゴリカル変数の個々のカテゴリをダミー変数化し、それに対して相関比指標を求めた上で、上記(3)の応用を考えることもできるが、0-1 データに対する相関比は概して（経験的・実務的に）大きくなりにくいので、より簡便には層化変数と目的変数であるカテゴリカル変数の間の連関係数を層化の効果の指標と代用する方法を考えた。本来の層化の効果を現す指標ではないが、実務的に扱いやすく、初期の検討に有効である。

##### (2) 層化に有効な変数の検討

この点についてはいくつかの研究をおこなった。

市区町村国政選挙投票率データ（おもに学会報告）

層化変数自体の性格を理解するための研究と位置づけられる。2004 年と 2007 年に行われた国政選挙での市区町村単位での投票率および政党得票率のデータを入手し、これに基づいて、従来型の人口規模で層化した場合の層化の効果を検討した。実際には、いくつかの標本設計（評価法 1：市区町村を単位とする層別無作為抽出、評価法 2：個人を単位とする層別無作為抽出、評価法 3：層別二段無作為抽出）の下での層化の効果や地点間の異質性を検討した結果から、例えば次のような知見が得られた。評価法 1 では、相関比による層化の効果を評価すると、市区町村投票率、自民党得票率などの変数で層化の効果が大きいこと。評価法 2 では、平成の大合併の効果で人口規模による層化では層間の異質性が（懸念されたとおり）低まったことを示唆する結果が得られたこと、評価法 3 では、層内での二段抽出時に考慮すべき地点間異質性は、都市規模が小さいほど大きくなる傾向が観察されること。

社会調査データを利用した層化変数の効果の検討（学会報告 など）

実務面での検討として実際の社会調査のデータを利用して、目的変数の性格の違いによる層化変数の効果の検討を行った。利用した調査データは統計数理研究所による「国民性に関する意識動向（2012 年度）調査」「日本人の国民性第 13 次全国調査」である。利用した層化変数は、人口集中地区の割合、高齢化率、大卒率などの地点の特徴を表す変数を 2010 年国勢調査から取得したもので、層化の単位としては市区町村レベルでの層化と、（可能な場合には）町丁字等レベルでの層化の結果とを、ベンチマークである「人口規模」による層化の場合と比較した。その結論として、目的変数によっては、大卒者の割合という教育水準に関わる変数が、人口集中

地区割合など人口密度や人口規模に関わる変数と同等かそれ以上の層別の効果を示す場合があり、有力な層化変数となり得ることを示唆する結果など、人口規模以外の変数を探索することの価値を示唆するものであった。なおこの結果は、関連研究である論文で、教育関連変数が多水準分析における地点間変動の有力な説明要因となるという知見と整合的であった。

### (3) 実証研究

平成 28 年度中に実施した Web 調査に基づく分析についてのみのべる。

この調査データについて対象者の居住市区町村を質問により判別し、居住市区町村の特性を表す変数を外部データより取得して層化を行い(5 層程度に分ける)、目的変数毎に層間の異質性に関わる指標を計算して、層化変数としての性能を比較するという分析を行うこととした。考慮する層化のための変数は、人口規模(これが従来からの層化基準であり、ベンチマーク)、国政レベルの選挙における市区町村投票率、2010 年国勢調査での大学進学率、などであった。

調査項目の回答は厳密に言えば量的変数と扱うべきではない順序尺度水準の変数が多かったが、議論の簡素化のために間隔尺度水準とみなし、平均値の層間の分離の程度を表す相関比指標で層化の効果を検討した。

結論的には考慮した層化変数は、平均的にはベンチマークである人口規模を上回る層化効果を必ずしも期待できず、従来からの方法が(多くの項目と強く連関するという意味で)無難だが有効な層化基準であることが再確認された。ただしこの結果が、目的変数によっては、人口規模よりも効果的な層化基準が存在する可能性を否定するものではない点を改めて強調すべきである。

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

#### [図書](計1件)

前田忠彦 鶴岡市民の方言と共通語の使い分けについて 第4回鶴岡市における言語調査の結果から 石川有香・石川慎一郎・清水裕子・田畑智司・長加奈子・前田忠彦(編)『言語研究と量的アプローチ』、金星堂、156-168, 2016年, 117-145.

#### [雑誌論文](計4件)

田中ゆかり, 林直樹, 前田忠彦, 相澤正夫 1万人調査からみた最新の方言・共通語意識 「2015年全国方言意識Web調査」の報告 査読あり 国語研論集 第11号, 2016年, 117-145. DOI:/10.10584/00000844

中村 隆・前田忠彦 日本人の国民性調査 第13次全国調査 よろん(日本世論調査協会報), 査読なし 115号, 2015年, 62-71.

Matsuoka, R. and Maeda, T. Neighborhood and Individual Factors Associated with Survey Response Behavior: A Multilevel Multinomial Regression Analysis of a Nationwide Survey in Japan, Social Science Japan Journal, 査読あり Vol.18, No.2 2016年, 217-232.

DOI:/10.1093/ssjj/jyv011

赤枝尚樹・前田忠彦 都市度尺度としての人口ポテンシャルの検討, 関西大学社会学部紀要 査読無し, 第45巻2号, 2014年, 249-265.

#### [学会発表](計11件)

前田忠彦 自治体投票率データの調査設計における利用に関する研究 2004年参院選を事例として 日本社会学会第89回大会(九州大学), 2016年10月

田中ゆかり, 林直樹, 前田忠彦, 相澤正夫 2015年全国方言意識Web調査に基づく話者累計 計量国語学会第60回記念大会(日本大学) 2016年10月

Maeda, T., Tanaka, Y., Hayashi, N. and Aizawa, M. Impacts of sociodemographic factors on the type of dialects usage in contemporary Japan, 3<sup>rd</sup> ISA Forum of Sociology (ウィーン大学), 2016年7月

Maeda, T., Inagaki, Y. Nakamura, T. and Park, Y-S. The Surveys on the Japanese National Character: Project history and some results from the most recent survey, 2015 Spring Meeting of the Korean Association for Survey Research, (Seoul, Republic of Korea), 2015年6月

稲垣佑典・中村 隆・前田忠彦・朴 堯星「国民性に関する意識動向調査 2012年度~14年度調査を通じてみられる意識・行動の変化と新たな知見」, 第60回数理社会学会大会(大阪経済大学), 2015年8月.

前田忠彦 社会調査における層化変数の効果に関する検討(その2), 第59回数理社会学会大会(久留米大学), 2015年3月.

前田忠彦・中村 隆 「日本人の国民性第13次全国調査」の設計と実施概要, 日本行動計量学会第42回大会(東北大学), 2014年9月

前田忠彦 社会調査における層化変数の効果に関する検討, 第58回数理社会学会大会(日本女子体育大学), 2014年8月.

中村 隆・前田忠彦・朴 堯星・尾崎幸謙 国民性に関する意識動向調査: 計画と実施概要, 日本行動計量学会第41回大会(東邦大学), 2013年9月

前田忠彦 社会調査における層化変数の選択, 第56回数理社会学会大会(関西学院大学), 2014年8月.

前田忠彦・大崎裕子 日本人の国民性調査における帰属階層項目の分析, 第56回数理社会学会大会(関西学院大学), 2014年8月.

6. 研究組織

(1) 研究代表者

前田 忠彦 (MAEDA, Tadahiko)  
統計数理研究所・データ科学研究系・准教授  
研究者番号：10247257

(2) 研究分担者 なし

(3) 連携研究者 なし

(4) 研究協力者 なし