

(案)

諮問第 57 号の答申

国民生活基礎調査に係る匿名データの作成について

本委員会は、諮問第 57 号による厚生労働省が作成を予定している平成 19 年国民生活基礎調査（以下「19 年調査」という。）に係る匿名データの作成方法の計画について審議した結果、下記の結論を得たので答申する。

記

1 計画の適否

本計画については、これにより作成される匿名データにおいて、19 年調査の調査客体の匿名性及び学術研究等における有用性がおおむね確保されるものと認められることから、適当である。

ただし、以下の「2 理由等」で指摘した事項については、修正が必要である。

2 理由等

(1) 19 年調査の匿名データ作成の基本的な方法

本計画では、平成 16 年国民生活基礎調査（以下「16 年調査」という。）の匿名データ作成方法に基本的に則り、19 年調査の匿名データを作成することとしている。

これについては、19 年調査で新たに追加された調査項目が少なく、かつ、それらについても 16 年調査の匿名データ作成方法に則って作成されること、また、16 年調査の匿名データ作成の際に見送った項目の一部が提供されること、さらに、社会情勢の変化を勘案して作成されることから、適当である。

(2) 匿名データ作成方法の変更

ア 年齢の上限値の変更

本計画では、世帯員の年齢について、一定の値を上限値とし、それを上回る場合に上限値以上でまとめる措置（以下、「トップコーディング」という。）を行うこととし、当該上限値を 90 歳以上としている。

これについては、前回の 16 年調査に係る答申（「諮問第 34 号の答申 国民生活基礎調査に係る匿名データの作成について」（平成 23 年 4 月 22 日府統委第 52 号））において、今後、匿名データの作成対象年次を拡大する際には、当該年次の年齢構成に応じて検討する必要がある旨、指摘されていることへの対応であり有用性が向上すること、また、平成 19 年の年齢構成を考慮すると匿名性が確保されていることから、適当である。

イ トップコーディングの上限値の変更

本計画では、1%閾値基準に基づいて、家計支出額についてトップコーディングの上限値を変更することとしている。具体的には、家計支出総額と子への仕送り額は上限値を低下させ、育児費用は上昇させることとしている。

これらについては、16年調査と同程度の匿名性を確保する必要から、前回設定した閾値に則ってデータを精査し、その結果上限値を変更することは適当である。上限値を低下させることとなる項目については、有用性を低下させることになるが、匿名性の確保の観点からやむを得ない措置である。ただし、育児費用については、有用性の観点とデータの分布状況から判断すると本計画の上限値7万円は7万5千円に変更することが必要である。

ウ 19年調査で新たに把握された調査項目の提供

本計画では、19年調査で新たに把握された調査項目について、16年調査の匿名データ作成方法に則り、匿名データ化することとしている。具体的には、世帯票の調査項目「すぐに仕事につけるか否か」は回答項目で提供、健康票の調査項目「こころの状態」は、出現頻度の低い項目を統合して提供することとしている。

これらについては、16年調査の匿名データ作成方法に則った措置であり、おおむね適当である。「こころの状態」については、有用性を懸念する指摘もあるが、今回、初めて提供されることを考慮するとやむを得ない措置である。

エ 16年調査で提供を見送った項目

本計画では、16年調査の匿名データ作成において提供を見送った項目の一部について、以下のように提供することとしている。

調査項目	19年調査での提供内容
手助けや見守りを要する者の状況 日常生活の自立の状況 自立期間 (主な介護者の状況) 手助けや見守りを要する者との続柄 同別居の別 性	回答項目で提供 出現頻度の低い項目を統合して提供 出現頻度の低い項目を統合して提供 回答項目で提供 回答項目で提供
自覚症状 症状名 最も気になる症状	回答項目で提供 回答項目で提供
通院している傷病名	回答項目で提供
普段の活動ができなかった日数 (16年調査は、「就床日数」として選択肢での回答形式)	実日数で提供することとし、25日以上をトップコーディング

これについては、16年調査で提供を見送った項目や調査項目の変更に伴う措置であり、有用性が向上し匿名性も確保されていることから、適当である。

オ トップコーディング等が行われた変数の基本統計量

本計画により作成された匿名データの各レコード上の変数のうち、トップコーディング等が行われた変数の基本統計量（平均値、中央値）等については、今後、利用者へのヒアリング等を行い、その結果に応じ、基本統計量について提供することとしている。

これについては、前回の答申における「今後の課題」への対応であり、適当である。

3 今後の課題

本計画においては、前回の16年調査に係る答申における指摘事項への対応が一部に留まっている。匿名データの利用者のニーズについては様々なものが考えられることから、以下の課題等について速やかに検討を進め、当該データのより一層の充実に努める必要がある。

(1) 地域情報の付与及び再抽出の単位

国民生活基礎調査に係る匿名データ作成においては、匿名性を確保するため、調査客体である世帯の特定につながる可能性が高い地域情報を削除し、地域区分を「全国」のみとする厳格な匿名化措置を講じることとしている。しかしながら、地域情報は、有用性の観点から極めて重要であることから、調査客体の匿名性の確保を十分に図りつつ、匿名データの利用者のニーズを踏まえて、何らかの地域情報を付与することの妥当性と可能性について検討する必要がある。

また、調査客体から再抽出（以下「リサンプリング」という。）する単位については、世帯単位のみとしているが、公衆衛生や疫学分野の研究においては、世帯員単位での健康状態や生活習慣の分析が重要となること等から、利用者のニーズを十分に考慮した上で、世帯員単位でのリサンプリングによる匿名データの作成の可能性について検討する必要がある。

(2) 所得票の内訳情報の提供

本計画では、16年調査同様、所得票に含まれる情報については、世帯の総所得、課税等の状況及び掛金のみ限定して提供することとしている。

しかしながら、近年、社会保障や所得格差等に関する研究の重要性が増しており、その分析には所得等に関する内訳や世帯員別の情報が重要であること、一方、本計画で適用されていないトップコーディング等以外の匿名化措置の適用も考えられることから、今後、匿名化措置に関する研究等の進展や利用者のニーズを十分に考慮した上で、所得等の内訳や世帯員別の情報の提供の妥当性と可能性について検討する必要がある。

(3) 匿名データの作成対象年次の拡大

本計画では、匿名データの作成対象調査を調査実施後5年以上経過したものとしており、今回は平成19年に実施したものを作成対象とするとともに、今後、順次拡大することとしている。

しかしながら、研究には経年的な分析が重要であるとともに、近年の経済・社会状況の急激な変化に伴い直近の統計に基づく分析の重要性が増していること、さらに、本調査については3年ごとに大規模調査が実施されていることを踏まえれば、提供時期の短縮について検討する必要がある。

また、年次の拡大に伴い提供項目の変更などの可能性も考えられることから、今後、利用者のニーズを踏まえて、匿名化措置の内容や組合せなどについて検討するとともに、内容等の変更があった場合には既に匿名データを作成・提供済みの年次調査への適用についても検討する必要がある。

諮問第57号「国民生活基礎調査に係る匿名データの作成について」

諮問文

諮問の概要

別添1 匿名データの作成方法の変更点

別添2 平成19年国民生活基礎調査に係る匿名データの作成方法（案）

別添3 平成19年国民生活基礎調査匿名データのチェックリスト

別添4 平成19年国民生活基礎調査の調査項目と匿名データの提供項目（対平成16年比較）

別添5 平成19年国民生活基礎調査の概要

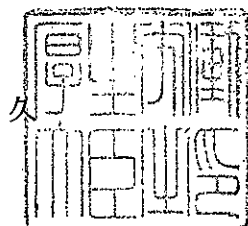
厚生労働省発統 0820 第 3 号

平成 25 年 8 月 20 日

統計委員会委員長

樋口 美雄 殿

厚生労働大臣
田村 憲久



諮問第 57 号

国民生活基礎調査に係る匿名データの作成について（諮問）

標記について、別紙のとおり作成するに当たり、統計法（平成 19 年法律第 53 号）
第 35 条第 2 項の規定に基づき、統計委員会の意見を求める。

諮 問 の 概 要

(平成 19 年国民生活基礎調査に係る匿名データの作成について)

今回、厚生労働省は、平成 19 年国民生活基礎調査（以下、「平成 19 年調査」という。）について、統計法（平成 19 年法律第 53 号）第 35 条第 1 項の規定に基づき、以下のとおり匿名データの作成を行う予定である。

1 平成 19 年調査の匿名データを作成する理由

本調査の匿名データにおいて、平成 16 年国民生活基礎調査（以下、「平成 16 年調査」という。）の匿名データ作成については、過去に答申（「諮問第 34 号の答申 国民生活基礎調査に係る匿名データの作成について」（平成 23 年 4 月 22 日府統委第 52 号））を得ており、また、平成 13 年国民生活基礎調査（以下、「平成 13 年調査」という。）の匿名データ作成についても、その答申内容に沿った匿名化手法に変更がないことから、統計委員会への諮問は不要とされ、現在、両年次のデータを提供しているところ、この度、平成 19 年調査についても同様の匿名化手法を用いつつ、社会情勢の変化や他調査の作成方法等を勘案し、匿名データを作成するものである。

2 匿名データの種類

平成 19 年調査の匿名データでは、平成 16 年調査及び平成 13 年調査と同様に、以下の 2 種類の匿名データを作成する。

(1) 匿名データ A（世帯票、健康票）：

人口、社会統計分野での世帯数の推計分析等を中心とした利用を想定

(2) 匿名データ B（世帯票、健康票、所得票及び貯蓄票）：

世帯の所得及び貯蓄に関する分析等を中心とした利用を想定

3 匿名データの作成方法の変更点

(1) 年齢のトップコーディング

個人の年齢階級の上限について、当該年の人口構成を踏まえ、「90 歳以上」に変更する。

(2) 1%閾値基準に基づく上限値（グルーピング）の変更

家計支出額、子への仕送り額及び育児費用のトップコーディングを変更する。

(3) 当該年次で把握された調査項目の提供

世帯票の「すぐに仕事につけるか否か」を提供するとともに、健康票の「こころの状態」について、出現頻度の低い項目を統合して提供する。

(4) 平成 16 年調査で提供を見送った項目への対応

世帯票の「手助けや見守りを要する者の状況」の提供を拡大するとともに、健康票の「自覚症状」及び「通院している傷病」を統合せずに提供する。（ただし、「最も気になる傷病」は従来通りの取り扱い。）

また、「就床日数」は「普段の活動ができなかった日数」と表現が変更になり実日数を記入することとなったため、実日数をトップコーディングする。