

第13回 匿名データ部会 議事概要

1 日 時 平成25年1月31日(木) 9:58~11:58

2 場 所 中央合同庁舎4号館2階 共用第3特別会議室

3 出席者

(部会長) 椿広計

(委員) 安部由起子、津谷典子

(専門委員) 伊藤伸介、加藤久和、安田聖

(審議協力者) 総務省(政策統括官(統計基準担当))、財務省、厚生労働省、農林水産省、経済産業省、国土交通省、東京都

(諮問者) 総務省統計局統計調査部: 井上調査企画課長、岩佐国勢統計課長ほか

(事務局) 内閣府統計委員会担当室: 若林参事官ほか

4 議 事

(1) 国勢調査に係る匿名データの作成について

(2) その他

5 議事概要

(1) 前回部会での質問に対する対応、及び残りの論点について

ア 年齢差の大きい親子のいる世帯の削除の状況について

前回部会において、有用性の観点から、年齢差が45歳以上の親と子のいる世帯について削除の状況を確認したいという意見があったことに対して、諮問者から、特に男親の場合に削除数が非常に多くなることから、親の性別による分布の差を考慮して、男親について年齢差を55歳以上とするという提案があった。各委員等からの主な意見は次のとおり。

- ・ 年齢差の大きい親子のいる世帯の削除数が多くなることを心配していた。この変数については、親の性別によって削除する年齢差を変える方法でよいと思う。

部会長のまとめ

- ・ 年齢差が大きい親子のいる世帯について、親の性別によって削除する年齢差を変えることとし、男親の場合に年齢差を55歳以上とすることは、これにより削除数を減らすことができるので、認めることとしたい。

イ 利用交通手段について

前回部会において、「利用交通手段」について、「オートバイ」と「自転車」は共に二輪で性質が似ていると考えられるため、これらを統合することはできないかという意見があったことに対して、諮問者から、「オートバイ」及び「自転車」を統合し、残りの「勤め先・学校のバス」、「ハイヤー・タクシー」及び「その他」を統合することとしたとしても、秘匿に関して特段問題は生じないという説明があった。各委員等からの主な意見は次のとおり。

- ・ 「オートバイ」及び「自転車」は性質が似ているので、秘匿上の問題がないのであれば、利用者の立場から見た場合に、これらを統合するのが望ましいと思う。

部会長のまとめ

- ・ 有用性の観点から、「オートバイ」及び「自転車」を統合し、残りの「勤め先・学校のバス」、「ハイヤー・タクシー」及び「その他」を統合することは、秘匿に関して特段問題は生じないと考えられるので、認めることとしたい。

ウ トップコーディングが行われた変数の基本統計量の提供について

前回部会において、トップコーディングが行われた変数の基本統計量の提供について地域ごとに提供できないかという意見があったことに対して、諮問者から、対象となる就業時間について地域ごとに提供するという説明が行われ、了承された。

エ 労働力状態について

- ・ 「労働力状態」については、「通学のかたわら仕事」を秘匿するため、「家事などのほか仕事」と統合することとしているが、平成 12 年調査で提供される調査項目と組み合わせた場合に、匿名性に問題は生じないか。

当初、「家事などのほか仕事」と「通学のかたわら仕事」を統合することを考えていたが、平成 12 年調査については、当該調査に含まれる調査項目との組み合わせによっては秘匿が不十分となる場合があることから、その調査項目自体を提供しないこととするか、就業者の内訳を提供しないこととするか、更に「休業者」を加えた3つを統合することとするか、あるいは、統合の仕方は変えず、外観識別性が高いと考えられるレコードを削除することとするかのいずれかにより対応したい。

- ・ 平成 12 年調査にある調査項目は有用性が高いこと、また削除数を最小限にとどめる方がよいということから、分類区分の統合か、あるいは就業者の内訳を提供しないこととする方がよい。
- ・ 「家事などのほか仕事」、「通学のかたわら仕事」に「休業者」を統合すれば問題は起こらないのではないか。「労働力状態」については自己申告なので、通学と仕事のどちらが主であるかもわかりにくいと考えられる。
- ・ 分類区分の統合による対応では、地域情報や他の情報と組み合わせた場合にまだ「通学のかたわら仕事」の秘匿が不十分となる可能性があるため、「通学のかたわら仕事」の出現頻度が低くなる地域を対象に外観識別性が高いと思われるレコードを削除する方がよい。
- ・ 外観識別性が高いと思われるレコードの削除という例外的な処理をするべきではなく、一定の方法で扱った方がよいのではないか。
- ・ 提供可能な調査項目を考えて、その上で 0.5%基準に照らして分類区分の統合を行い、それでも秘匿が難しいという場合にレコードの削除を行うということであり、その意味では、例外にはならないのではないかと考える。
- ・ 更に「休業者」を統合するだけでは秘匿性の観点でリスクがあるということであれば、有用性は劣るものの、就業者の内訳を提供しないとするのがよいのではないか。
- ・ 今回は就業者の内訳を提供しないことに賛成するが、今後検討の余地がある。

部会長のまとめ

- ・ 秘匿を重視する観点から、平成 12 年調査については就業者の内訳を提供しないということでもとめることとしたい。

オ 「スワッピング」の説明について

- ・ 「スワッピング」については、その具体的な方法を一般に説明することができないことから、例えば、スワッピングの前後で影響を受ける項目の分布にあまり差異がないというようなことを、ユ

ーザがわかるような形で説明するということが考えられるが、総務省統計局ではどのようなことを考えているか。

スワッピングを行っていること、及びそれによる誤差は標本誤差と比較して許容範囲となっており有用性に支障はないことについて、ホームページ等で説明することを考えている。また、主要な項目の全国の結果について、母集団と匿名データとを比較した場合にそれほど乖離がないという情報についても、併せて提供したいと考えている。

- ・ 母集団と匿名データとの比較の情報について、全国のみ提供するということが、特定の地域のデータを分析したいと考えている利用者にとっては、それでは情報が足りないのではないか。

利用者の関心のある地域に応じてそのようなデータを個別に提供することは困難であり、提供するデータの量も多くなってしまふ。地域別の結果の比較については、匿名データによる集計結果と詳細な地域の公表統計により行うことができる。

- ・ 国勢調査では地域別の詳細な報告書が公表されているので、スワッピングで基本的な属性に歪みは生じないということを示すという点では、全国の結果の比較だけを提供すればよいのではないか。
- ・ 学会や学術誌での発表などを行う際に、どの程度の詳細な地域までの分析が保証されるのかが、利用者が一番の関心事項であると思うが、今回の説明では、その点がわからない。例えば、都道府県の範囲の分析であれば保証されている、あるいは都道府県内でスワッピングを行っているといったようなことを示すことはできないか。
- ・ 地域での分析が保証されているということを説明できれば、査読付き学術誌への投稿論文の審査等などの関係で、利用価値は非常に大きくなると考えられる。
- ・ 利用者がスワッピングについて詳しいとは限らないので、諸外国の人口センサスの匿名化の際にも用いられている手法であるということなどを、ウェブサイト等に掲載するという対応をしていただければよいと思う。
- ・ 利用者への説明については、書きぶりにもよるが、スワッピングが都道府県内で行われていることを開示することのリスクは、基本的にあまりないと考えられる。

部会長のまとめ

- ・ 都道府県内では分析に影響を与えないということを説明することで、研究の有用性が格段に向上し、そのことによるリスクはそれほど増えていないと考えられるので、スワッピングが都道府県内で行われているということを開示すべきである。

(2) 答申(案)の審議について

椿部会長から、資料「諮問第44号の答申『国勢調査に係る匿名データの作成について』(案)」が示され、事務局からの答申(案)についての説明の後、審議が行われた結果、一部で所要の修正を行うこととされたが、答申(案)はおおむね適当であるとして採択された。なお、答申(案)の修正部分の表現については、部会長に一任することとされた。各委員等からの主な意見は次のとおり。

ア 出現頻度が低い又は特徴的な値があるレコードを含む世帯の削除(2の(1)のイ)

- ・ 本日の議論で変更のあった年齢差の大きい親子のいる世帯の削除については、特に父子について削除数が非常に多くなっていることから、年齢差が55歳以上の男親と子と修正する必要がある。

イ 労働力状態(2の(2)のアの(ウ))

- ・ 平成12年調査に関しては、秘匿が不十分となることから、就業者の内訳を提供しないようにする必要があるという修正を行うこととする。

ウ 利用交通手段（２の（２）のアの（カ））

- ・ 「オートバイ」と「自転車」を統合するというので、答申案の案２を採用することとする。

エ スワッピング（２の（３）のア）

- ・ スワッピングの導入について、「やむを得ない」と記述している部分については、国勢調査は悉皆調査であり秘匿性の観点からスワッピングを行う必要があること、及び諸外国の人口センサスでも導入していることから、はっきりと「必要である」と表現とした方がよい。
- ・ スワッピングの導入が都道府県の分析に影響を与えないことについては、そのことを開示しても調査客体の特定化の危険性が低いということを明記した方がよい。

部会長のまとめ

- ・ スワッピングの導入について、「やむを得ない」となっている部分については、「必要である」と修正したい。
- ・ スワッピングが行われていることがわかっている以上、特定化のリスクはあまり増えないということを中心に、スワッピングは都道府県内で行われているということを開示するように修正したい。

オ トップコーディングを行う高齢者の年齢の検討（３の（１））

- ・ トップコーディングを行う高齢者の年齢の検討に関して、今後の作成対象年次の拡大の際に「人口構成」に応じた検討が必要と記述している部分については、どのような人口の構成かはっきりとわかるように、「人口の年齢構成」とした方がよい。

部会長のまとめ

- ・ 「人口構成」という表現については、「人口の年齢構成」に修正することとしたい。

カ 複数の匿名データの作成の可能性に関する検討（３の（２））

- ・ 「提供データを個人単位とすることなどにより」の「など」の中には、地域区分をより詳細にするという匿名化を緩和する方向だけでなく、場合によっては詳細な地域情報の提供を可能にするために、ある特定の調査項目を提供しないといった、匿名化を厳しくする方向も含まれていると考えた方がよい。
- ・ 匿名化措置については、その技法を中心にこれまで議論されているが、諸外国では、例えば学術目的でより詳細なデータを提供するために、法制度の観点から匿名データの提供方法を変えているところもあり、将来的には、そのような観点からの匿名データの提供の可能性もあり得るのではないか。

（３）その他

- ・ 本日は承された答申（案）及び議事概要については、２月１５日（金）の統計委員会に提出することとされた。

以上

<文責 内閣府大臣官房統計委員会担当室 速報のため事後修正の可能性あり>