

第 24 回 匿名データ部会 議事概要

- 1 日 時 平成29年11月29日（水） 17:00～19:00
- 2 場 所 総務省第二庁舎 6階 特別会議室
- 3 出席者
- （部 会 長） 北村 行伸
- （委 員） 川崎 茂、永瀬 伸子
- （専 門 委 員） 神林 龍、寺村絵里子、南 和宏
- （審議協力者） 総務省（政策統括官（統計基準担当））、総務省統計局、財務省、文部科学省、厚生労働省、経済産業省、国土交通省、東京都、千葉県
- （諮 問 者） 総務省：統計局統計調査部調査企画課 栗田課長ほか
厚生労働省：政策統括官（統計・情報政策担当）付参事官付審査解析室 田中室長ほか
- （事 務 局） 総務省：横山大臣官房審議官
総務省統計委員会担当室：上田次長ほか

4 議 事

- （1）住宅・土地統計調査に係る匿名データの作成について
- （2）国民生活基礎調査に係る匿名データの作成について
- （3）その他

5 議事概要

（1）住宅・土地統計調査に係る匿名データの作成について

冒頭、現在検討されている次期基本計画における匿名データを含めた調査票情報の提供の検討状況について、事務局と総務省統計局から報告があった。

また、総務省統計局から、複数の匿名データを作成することについて、次期基本計画を先取りして、統計研究研修所の支援を受けて、具体的な検討を開始している旨、説明があった。

その後、答申案について審議が行われ、一部文言等の修正を行うことを前提に了承された。答申案については、所要の修正を行った後、第 117 回統計委員会（平成 29 年12月 19 日開催予定）において部会長から報告することとされた。

答申案に係る主な意見は以下のとおり。

- 新規追加項目に関する最終行の「提供しないことは適当である」の記述は、ネガティブな印象を与えるので、何か表現を工夫すべきではないか。
- 検討したい。

○複数の匿名データ作成に関して、「匿名レベル」は、あまり聞きなれた言葉ではないので、表現を変えるようなことが必要ではないか。また、「国勢調査の匿名データを利用した」とあるが、国勢調査の匿名データを利用して検討するのではなく、国勢調査の匿名データを作成するための検討プロセスの中で得た情報などを利用して検討するということだとすれば、言葉が足りない気がする。

→検討したい。

○複数の匿名データ作成に関して、「より詳細な地域情報」とあるが、都道府県レベルより詳細などと比較するレベルを書く必要がある。

→検討したい。

(2) 国民生活基礎調査に係る匿名データの作成について

厚生労働省から諮問の概要について説明の後、論点メモに沿って審議が行われ、次回、引き続き審議することとされた。

主な意見は以下のとおり。

○リサンプリング方法については、見直してもいい時期にきているのではないか。ある世帯が調査対象になったら、その近隣世帯の情報が分かりやすくなるという説明は、根拠が薄弱である。地域情報を提供しないことを前提にすれば、ダイレクトサンプリングでも構わないと思う。今回の作業で間に合わなければ、次回までに検討してほしい。

○抽出された集落は悉皆だというのが、リサンプリングされた結果、全体の2割しか抽出されていないのであれば、全国全体から抽出してもよいと思う。

○今後、リサンプリングを見直すのであれば、抽出率の8割や標本一意であったときに母集団でどのくらいの確率で一意になっているのかを評価するとよいのではないか。

○匿名データCを作成して他の匿名データとマッチングすると、かなりのデータが特定される可能性があるとの説明だが、同じ世帯が選ばれているということが認識できる可能性はあるが、どこの世帯かという秘匿性は破られてはいないので、その説明は誤解を招きやすい。

○本体調査が二段階抽出なので、匿名データも二段階抽出をしているが、効率性はよくないし、むしろ特定されるリスクが高いのではないかと考えるので、検討に時間を要するのであれば、引き続きの課題にしてもよいと思う。

○外観識別性の高い属性は何かを定義し、その属性で度数分布表を作成し、そこに地域情報を入れる、入れないで、一意識別がどう増えるかを検証すれば、秘匿性の確保が定量的に分析できると思う。

○次回に説明される際に、第一段階のリサンプリング率のマックス8割と第二段階の一律8割について、どうして固定なのか説明してほしい。

○都道府県よりも粗い地方ブロックのような地域情報は、医療を専門に研究される方にとっては使えない情報とのことだが、使えなくてもリサンプリングには役に立つかもしれない。

- 所得の内訳を提供していただくことは有難い。欲を言えば、世帯主と配偶者の内訳が分かると、もっとよい。
- 所得の内訳で、雇用者所得と公的年金・恩給以外の項目を提供しない理由は何か。また、世帯の総所得は提供せずに、世帯員の所得だけを提供するという選択肢はないのか。
- 世帯員の情報を提供することは、リサンプリングの方法に関連するが、世帯員でのリサンプリングは困難と考えている。

(部会長のまとめ)

- 新規に追加された事項を提供すること、所得の内訳情報として「雇用者所得」と「公的年金・恩給」を提供することは適当と考える。ただし、リサンプリングの方法とそれに関連して地域情報の付与については、次回、もう一度議論したい。

以上

<文責 総務省統計委員会担当室 速報のため事後修正の可能性あり>