

諮問第 13 号の答申

全国消費実態調査、社会生活基本調査、就業構造基本調査及び住宅・土地統計調査に係る匿名データの作成について（案）

本委員会は、総務省が平成 21 年に作成を予定している全国消費実態調査（指定統計第 97 号を作成するための調査）、社会生活基本調査（指定統計第 114 号を作成するための調査）、就業構造基本調査（指定統計第 87 号を作成するための調査）及び住宅・土地統計調査（指定統計第 14 号を作成するための調査）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 計画の適否とその理由等

(1) 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、全国消費実態調査等 4 調査（以下「作成対象 4 調査」という。）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、一部修正を行うことを前提に適当なものとする。

この判断の理由及び修正点は以下のとおりである。

(2) 理由及び修正点

ア 情報の削除

(ア) レコードのリサンプリング

匿名データの作成に当たっては、作成対象 4 調査の全ての標本のレコード（調査客体）から、世帯単位により、全国消費実態調査、社会生活基本調査及び就業構造基本調査の 3 調査については 80%を、また、住宅・土地統計調査については 10%を、無作為または各レコードに付された乗率の大きさに基づく確率比例で再抽出（以下「リサンプリング」という。）したもの（以下「サブサンプル」という。）を用いる計画である。

これについては、次の理由等から適当である。

リサンプリングは、匿名データの中に特定の調査客体が含まれるか否かの判別を困難とする措置であること

特に、今回のリサンプリングにおいては、無作為抽出を基本としつつ、各レコードが持つ集計用乗率に抽出地域との一定の対応関係がある場合、当該乗率から抽出地域が特定されてしまうことを防ぐための措置を採っていること

世帯単位による抽出は、匿名データの利用者のニーズが高い世帯収支等世帯に着目した分析が可能となるため、個人単位による抽出よりも当該データの有用性が高まること

サブサンプルの抽出率は、各調査の母集団の大きさやそれに含まれる情報の内容等を踏まえ設定しているものであり、当該抽出率によりリサンプリングされたサブサンプルから作成された匿名データによる統計と全レコードから作成された公表統計（以下「公表統計」という。）との間で、代表的な項

目の平均値や標準偏差に大きな乖離はなく、当該データの有用性が確保されていること

(イ) 識別情報の削除等

作成対象 4 調査のサブサンプル中のレコードに含まれる情報のうち調査区番号等の識別情報は、これを削除するとともに、当該レコードを乱数により並び替える計画である。

これらについては、調査客体の特定や探索を防止するために効果的な措置であること等から、適当である。

(ウ) 裾切りによるレコード削除

a 世帯人員 8 人以上等の世帯

作成対象 4 調査のサブサンプル中のレコードのうち世帯人員 8 人以上の世帯及び三つ子以上のいる世帯に係るものは、匿名データから削除する計画である。

これについては、世帯員の人数等の情報は世帯の外部から比較的容易に把握可能な属性であり、それが極端に大きい場合は調査客体が特定される可能性が生じること等から、適当である。

b 年収等が高額な世帯

全国消費実態調査のサブサンプル中のレコードのうち年収、貯蓄及び借入金がある一定金額以上の高額な世帯に係るものは、匿名データから削除する計画である。

これについては、匿名データの信頼性の確保の観点からサブサンプルの削除は必要最小限に留めるべきであること、年収等の情報は世帯外からの把握可能性が低いこと、世帯収支等の経済分析に対する研究者のニーズが非常に高いこと等から、年収等が高額な世帯のレコードを全面的に削除することは適当でない。したがって、当該世帯のレコードについても、提供する情報を年収等の総額のみに限定し、かつ、年収等の総額が高額な世帯については、一定の水準を上限値とし、こ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。）等の匿名化措置を講じた上で、匿名データに残すことが必要である。

イ 識別情報の階級区分の統合

(ア) トップコーディング及びボトムコーディング

a 高齢者の年齢

4 調査の匿名データの各レコード上の個人の年齢について、一定年齢を上限値とし、それを上回る高齢者の場合、トップコーディングを行うこととし、当該上限値は、全国消費実態調査及び住宅・土地統計調査は 75 歳以上、社会生活基本調査は 85 歳以上、就業構造基本調査は 80 歳以上とする計画である。

これについては、トップコーディングは、それにより極めて高齢であるという特殊な属性をまとめられ、調査客体の判別を困難とすることから適当であるが、上限値については、近年の高齢化の進展状況、年齢を用いた就業行動や家族関係の分析の重要性等を踏まえ、4 調査とも 85 歳以上にすることにより、匿名データの有用性の向上を図ることが必要である。

b 住宅の規模等

全国消費実態調査及び住宅・土地統計調査の匿名データの各レコード上の住宅の規模等に係る数値（延べ床面積、敷地面積、家賃・間代等）について、トップコーディング及び一定の水準を下限值としこれを下回る場合に下限値以下でまとめる措置（以下「ボトムコーディング」という。）を講じる計画である。

これについては、トップコーディング及びボトムコーディングにより、住宅の規模等が極端に大きい（または小さい）という特殊な属性をまとめられ、住宅の特徴を通じそこに居住する調査客体の判別を困難とすること等から、適当である。

(イ) リコーディング（分類区分の再付与）

a 地理的情報（地域区分）

匿名データの各レコードに付与する地理的情報については、その分類の程度を粗いものとする措置（以下「リコーディング」という。）を講じることとし、住宅・土地統計調査では 47 都道府県別に、また、全国消費実態調査、社会生活基本調査及び就業構造基本調査では全国 6 ブロック別とする計画である。

このうち、住宅・土地統計調査の地域区分については、当該調査の標本規模が非常に大きく、かつサブサンプルの抽出率が 10%と低いこと等により、地域区分と他の情報との組み合わせにより調査客体が特定される可能性が極めて低いこと等から、適当である。

一方、全国消費実態調査等 3 調査の地域区分については、3 調査のサブサンプルの抽出率が 80%と高いこと、3 調査は調査客体である個人の職業、配偶者との年齢差等多くの属性情報を有しており、これと詳細な地理的情報を組み合わせると調査客体の特定の可能性が生じること、地域区分を 6 ブロックとしても、公表統計と照合することにより都道府県の別が明らかになるケースが一部あること等から、地域区分を「3 大都市圏」及び「その他の地域」の 2 区分とすることにより、調査客体の匿名性の確保を十分に図るよう、万全を期すことが必要である。

b 個人の年齢

4 調査の匿名データの各レコード上の個人（トップコーディングを行う高齢者を除く。）の年齢については、15 歳以上の者は、リコーディングとして、5 歳階級別とする一方、15 歳未満の者は、リコーディングを行わず各歳別とする計画である。

このうち、15 歳以上の者については、各歳別のデータ提供に比べ、匿名データの有用性が低下するものの、各歳別の年齢が明らかになると、個人の職業等他の属性情報との組み合わせにより調査客体が特定される可能性が生じることから、やむを得ない措置である。

これに対して、15 歳未満の者については、調査客体の特定に利用可能な属性情報が限定されているため、各歳別の年齢を明らかにしても判別を困難とする観点からは適当である。

2 今後の課題

本計画については、政府における匿名データの作成は今回の総務省によるものが初めてであり、調査客体の匿名性の確保に慎重を期する必要があること、本年4月の統計法の全面施行に合わせて、匿名データの提供を速やかに開始する必要があること、これまで匿名データの利用ニーズが必ずしも十分に把握されていないこと等から、調査客体の匿名性を確保するために厳格な匿名化措置を講じていることはやむを得ない。

しかしながら、匿名データの利用者ニーズ等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) 本計画では、匿名性を確保するため、個人の年齢等調査客体の特定につながる可能性がある重要かつ基本的な属性情報については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、調査客体の匿名性は、一つの匿名化措置のみで確保される訳ではなく、複数の匿名化措置により全体として確保されるものであるため、匿名化措置の内容や組合せを変えることにより、同一の調査について複数の匿名データを作成することが可能であると考えられる。例えば、就業構造基本調査について、15歳以上の世帯員の年齢を各歳別とする一方、職業、産業等の分類区分を大括り化した匿名データの作成についてのニーズも指摘されている。
こうした観点から、今後、複数の匿名データのマッチングによる調査客体の特定の危険性に関する研究等の結果や匿名データの利用者のニーズを踏まえて、匿名化措置を課す情報及びその程度が異なる複数の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) 本計画では、匿名データの作成対象調査を平成元年以降に実施したものであり、かつ調査実施後5年以上を経過したものである。
しかし、経済・社会事象に関する研究には、長期の時系列分析が不可欠であり、また、近年、経済・社会の状況がめまぐるしく変化していることから、直近の統計に基づく当該研究の重要性も増している。
こうした観点から、今後、作成対象調査を、平成元年より前に実施したものに拡張することについて検討するとともに調査実施後5年以上経過したものを提供するという基準を緩和することについて検討する必要がある。
- (3) 匿名データの分析手法としては、集計値の分析のほかに多変量解析がある。しかし、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、ボトムコーディング及びリコーディング（年齢等の階級化等）が行われている変数については、集計値の分析には大きな問題がないものの、多変量解析へは必ずしも十分利用することができない。
こうした観点から、今後、トップコーディング等が行われた変数についても多変量解析に十分利用できるよう、当該変数の平均値等をメタデータとして整備する等の措置に関して、運用後のニーズ等の状況を踏まえ、検討する必要がある。



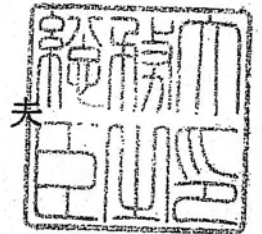
総統調第372号

平成20年12月22日

統計委員会委員長

竹内 啓 殿

総務大臣
鳩山 邦 夫



諮問第13号

全国消費実態調査、社会生活基本調査、就業構造基本調査

及び住宅・土地統計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項及び附則第3条の規定に基づき、統計委員会の意見を求める。

諮問の概要

1 匿名データの作成の対象とする統計調査

今回、総務省は、以下に掲げる統計調査について、統計法（平成 19 年法律第 53 号）第 35 条第 1 項の規定に基づき匿名データの作成を行う予定である。

匿名データを作成する統計調査名	調査年次
全国消費実態調査	平成元、6、11、16 年
社会生活基本調査	平成 3、8、13 年
就業構造基本調査	平成 4、9、14 年
住宅・土地統計調査	平成 5、10、15 年

（説明）

一般に、世帯・個人は事業所・企業よりも特性のばらつきが小さく、外部情報との照合により特定される可能性が低いため、世帯・個人を対象とする統計調査の方が、事業所・企業を対象とするものよりも、識別可能性のリスクの観点からは匿名データの作成は容易であるとされている。一方、世帯・個人を対象とする統計調査であっても、同一の調査客体を複数回継続的に調査するものや悉皆調査については、匿名データの作成が比較的困難であるとされている。

総務省では、このような点を踏まえ、まず、世帯・個人を対象とする上に掲げた統計調査について匿名データを作成することとした。なお、これらの統計調査については、総務省が一橋大学と共同で行ってきた「学術研究のための政府統計マイクロデータの試行的提供」（平成 16～20 年）の研究において、調査票情報に秘匿措置を講じた場合の当該データの安全性、有用性等について、研究してきたものである。

2 匿名データの作成方法の概要

上記 1 に掲げる統計調査について、匿名化措置を講じ、匿名データを作成することとし、その概要については以下のとおりである。

- ・ 元の統計調査のレコードすべてを匿名データに用いるのではなく、それに間引きを施したものをを用いる（レコードのリサンプリング）。
- ・ 識別情報は、レコードから全面的に削除する。また、レコードの配列順が意味をなさないように、無作為に並べ替えを行う（識別情報の削除等）。
- ・ 特徴的な識別情報の値があるレコードは、削除する（裾切りによるレコード削除）。
- ・ 極端に大きな値は、上限値を設けて頭打ちにする（トップコーディング）。
- ・ 分類事項の程度は、詳細なものではなく、粗いものとする（リコーディング）。

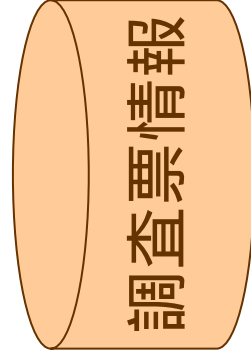
匿名データの作成方法の概要

(1) 情報の削除

- ア レコードのリサンプリング 元の統計調査のレコードすべてを匿名データに用いるのではなく、それに間引きを施したものをを用いる。
- イ 識別情報の削除等 識別情報は、レコードから全面的に削除する。
また、レコードの配列順が意味をなさないように、無作為に並べ替えを行う。
- ウ 裾切りによるレコード削除 特徴的な識別情報の値があるレコードは、削除する

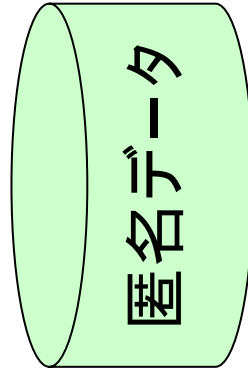
(2) 識別情報の階級区分統合

- ア トップ（ボトム）コーディング 極端に大きな値は、上限値を設けて頭打ちにする
- イ リコーディング 分類事項の程度は、詳細なものではなく、粗いものとする



匿名化

-  全国消費実態調査
-  社会生活基本調査
-  就業構造基本調査
-  住宅・土地統計調査



匿名データ部会の審議状況について（報告）

< 全国消費実態調査等関連 >

第2回匿名データ部会 議事概要

1 日 時 平成21年2月13日(金) 10:00~12:00

2 場 所 中央合同庁舎第4号館2階 共用第3特別会議室

3 出席者

廣松部会長、井伊部会長代理、宇賀臨時委員、椿臨時委員、津谷臨時委員、玄田専門委員、西郷専門委員、永瀬専門委員、星野専門委員、安田専門委員

総務省(政策統括官室(統計基準担当))、厚生労働省、経済産業省、国土交通省、東京都、千葉県、日本銀行

【諮問者】

杉山総務省統計局統計調査部調査企画課長、栗原総務省統計局統計調査部調査企画課調査官

【事務局】

中島内閣府大臣官房統計委員会担当室長、高木内閣府大臣官房統計委員会担当室参事官

4 議事次第 (1) 全国消費実態調査、社会生活基本調査、就業構造基本調査及び住宅・土地統計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) 全国消費実態調査、社会生活基本調査、就業構造基本調査及び住宅・土地統計調査に係る匿名データの作成について

「配付資料」の部会長論点メモにおける個別の論点について、以下のとおり、諮問者である総務省統計局から説明を受けた後に審議が行われた。

ア リサンプリングの方法

《総務省統計局の説明》

リサンプリング率については、作成計画案の抽出率により仮作成した匿名データによる統計と全標本から作成している公表統計との間で、代表的な項目に関し平均値等を比較した結果、乖離が少ないことが確認できたこと等から、適当であると考えます。

世帯単位によるサブサンプルの抽出については、過去の研究においても世帯関係に着目した分析が多く、利用者ニーズが高いと考えられることから、個人単位によるサブサンプルの抽出よりも適当なものであると考えます。

《各委員の主な意見》

- ・ 世帯人員8人以上の世帯のレコードについては、匿名性を確保するため、匿名データから削除することとしているが、個人単位によるサブサンプルの抽出を導入することにより、一部でもデータを残すことを検討すべきではないか。

- ・ 過去の一橋大学における研究例から見ると、世帯人員 8 人以上の世帯について、個人単位によるサブサンプルの抽出を導入しても、個人単位のみ情報は利用されない可能性が高く、当該導入の必要性は乏しい。
- ・ 標本の各レコードが持つ集計用乗率と抽出地域に一定の対応関係がある場合、当該乗率から抽出地域が特定されることを防ぐため、レコードを乗率別に層化した上で各層から乗率の大きさに応じて確率比例抽出を行い、その層の平均乗数を再付与する措置を講じているが、層化していれば層内の乗率にあまり違いがないことから、確率比例抽出まで行う必要はないのではないか。
- ・ リサンプリング率及びサブサンプルの抽出単位については、匿名性及び有用性の確保の観点から見て、作成計画案の考え方で問題はない。

イ 地域区分

《総務省統計局の説明》

地域区分を全国 6 ブロック別とした「全国消費実態調査」等 3 調査については、前回部会の指摘を受けて検討した結果、公表統計との照合により都道府県の別が明らかになる事例が一部あることが判明したため、「3 大都市圏」及び「その他の地域」の 2 区分にすることとしたい。

一方、住宅・土地統計調査の場合、世帯・個人等に関する情報が相対的に少ないことを勘案し、都道府県別とする。

《各委員の主な意見》

- ・ 3 調査について 2 区分とすることは保守的であると思うものの、匿名性の確保の重要性を勘案すると、今回はやむを得ない。
- ・ 3 調査における地域区分の細分化は、将来の検討課題である。

ウ 個人の年齢

《総務省統計局の説明》

年齢については、調査客体の特定リスクが大きいため、諸外国の例も勘案し、15 歳以上の者は 5 歳階級別としたい。

年齢に係るトップコーディングの上限値については、前回部会での指摘を受け、4 調査とも一律 85 歳以上としたい。

15 歳未満の若年者については、調査客体の特定に利用できる属性情報が限定されていることから、各歳別に提供することとする。

《各委員の主な意見》

- ・ 就業構造基本調査については、年齢を各歳別にしても、調査客体が特定されるリスクは少ないのではないかと。労働分野の研究においては、詳細な職業分類に関する研究よりも年齢を中心とした研究の方が主流なので、職業分類を 10 分類程度とする一方で、年齢を各歳別に提供して欲しい。
- ・ 労働分野の研究実績から見て、就業構造基本調査の匿名データにおいて各歳別に提供する必要性は低いのではないかと。
- ・ 年齢を各歳別にする一方、職業等の分類区分を大幅に粗くしたデータセットをもう一つ作ることはできないかと。

- ・ データセットを2つ作ると、これらの結合によってリサンプリングの匿名化効果が失われる危険性が生じ、安全性が低下する。
- ・ 15歳以上の者の年齢区分を5歳階級別とすることについては、匿名データの作成が今回初めてであることや調査客体の匿名性の確保の重要性等を勘案すると、やむを得ない。
- ・ 現時点では、年齢に係るトップコーディングの上限値を4調査とも一律85歳以上とすることは適当だが、近年の高齢化の進展状況等を勘案すると、将来、当該上限値の見直しを行う必要がある。

エ 世帯人員、床面積、家賃等

《総務省統計局の説明》

世帯人員が8人以上の世帯については、当該世帯レコードの中の他の情報との関係からトップコーディングが困難であり、当該世帯レコードそのものを削除することが必要である。

床面積等については、地域間の比較を考えた場合、都道府県毎にばらばらの上限値を設定しても有効な比較が出来ないこと等から、一律にトップ（ボトム）コーディングの上（下）限値を設定する。

トップコーディングを行った場合の平均値等の提供については、運用後のニーズ等の状況を踏まえ、検討することとしたい。

《各委員の主な意見》

- ・ 世帯人員が8人以上の世帯を削除することは適当だが、近年の世帯規模の縮小傾向等を勘案すると、将来、削除すべき世帯の世帯員数について見直しを行う必要がある。
- ・ 家賃等のトップ（ボトム）コーディングの上限値が全国一律でないとは有効な地域比較ができないとする説明には疑問がある。
- ・ 家賃等のトップコーディングの上限値を9万円とすることは、都市経済学分野で大都市圏の家賃分析を行っている研究者にとっては、匿名データの有用性の低下につながるのではないか。
- ・ 床面積、家賃等におけるトップ（ボトム）コーディングの上（下）限値については、匿名データの作成が今回初めてであること等を勘案すると、やむを得ない。

オ 全国消費実態調査における年収等が高額な世帯

《総務省統計局の説明》

当初、年収等が高額な世帯は匿名データから削除することとしていたが、前回部会での議論を踏まえ、その取扱いを再検討した結果、当初案どおり削除（案1） 0.5%基準に従って削除（案2の1） 提供する項目を限定し0.5%基準によるトップコーディングを実施（案2の2） 資産関係項目を活かし関連する他の情報を削除（案3） の4つの見直し案のいずれかにすることとしたい。

《各委員の主な意見》

- ・ 匿名データの信頼性の確保の観点から、サブサンプルの削除は必要最小限に留めるべきであること等を総合的に勘案し、案2の2が適当である。
- ・ 全国消費実態調査は、年収等に関する具体的な数字が出ており、また、他の統計調査以上に

調査客体との信頼関係で成り立っている調査なので、匿名性の確保を重視すべき。

カ 匿名データの作成対象年次

《総務省統計局の説明》

提供する匿名データについては、調査客体の匿名性の確保の観点から、平成以降の調査で、かつ調査実施から5年以上経過したものとしているが、今後、平成より前に実施したものも提供できるように検討したい。

《各委員の主な意見》

- ・ 平成より前に実施した調査のみならず、直近の調査についても今後検討すべき課題である。

キ その他（匿名化措置の考え方等）

《総務省統計局の説明》

0.5%基準は、諸外国においても、トップコーディングに関する基準として適用しており、全ての情報について適用している訳ではない。

前回部会において学生等のためにハードルを下げたデータを作るよりも訓練する機会等を提供することが重要といった意見があったことや、レコード同士の照合により安全面での問題があるとの意見もあったことから、複数の匿名データを作成することは適切ではないと考える。

《各委員の主な意見》

- ・ 全ての変数に0.5%基準を適用する必要はないが、極めて珍しい職業や外から見て分かる変数については、調査客体の特定リスクの有無を確認した方が良い。
- ・ 最終的に匿名データを提供する段階で、該当する個体数が極端に少ないものの有無を確認し、その結果、調査客体が特定される危険がある場合には、必要な処置を行うべきである。

(2) その他

今回の匿名データ部会は2月24日（火）に開催することとなった。

<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>

第3回匿名データ部会 議事概要

1 日 時 平成21年2月24日(火) 16:00~17:50

2 場 所 中央合同庁舎第4号館2階 共用第3特別会議室

3 出席者

廣松部会長、井伊部会長代理、宇賀臨時委員、津谷臨時委員、玄田専門委員、西郷専門委員、星野専門委員、安田専門委員

総務省(政策統括官室(統計基準担当))、厚生労働省、農林水産省、経済産業省、国土交通省、東京都、千葉県、日本銀行

【諮問者】

杉山総務省統計局統計調査部調査企画課長、栗原総務省統計局統計調査部調査企画課調査官

【事務局】

中島内閣府大臣官房統計委員会担当室長、高木内閣府大臣官房統計委員会担当室参事官

4 議事次第 (1) 全国消費実態調査、社会生活基本調査、就業構造基本調査及び住宅・土地統計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) 全国消費実態調査、社会生活基本調査、就業構造基本調査及び住宅・土地統計調査に係る匿名データの作成について

廣松部会長から、これまでの部会審議を基に作成された「答申(案)」が示され、事務局による朗読の後、項目ごとに順次審議が行われた。

各委員等の主な意見は次のとおり。

【計画の適否とその理由等】

裾きりによるレコード削除(年収等が高額な世帯:答申案1-(2)-ア-(ウ)-b)

- ・ 年収等が高額な世帯に対する匿名化措置に関し「全体に占める構成比が0.5%未満の変数」との記載があるが、ここでは世帯年収の総額に関することのみを記述しており、当該記載では多くの変数に関する記述のように誤解される恐れがあるため、表現を修正すべき。

【今後の課題】

ア 前文

- ・ 今後の課題に挙げられた点については、いずれも早急に検討すべき課題であることから、前文に「速やかに検討」という記述を追加すべき。
- ・ 匿名化措置の基準については、今後、利用者ニーズ等に応じて変更されるよう定期的に見直す機会を設ける旨の記述を追加する必要があるのではないか。

- ・ 匿名化措置の基準に関する定期的な見直しについては、前文の「以下の課題等について検討を進め」の部分に含意されている。

イ 同一の調査に係る複数の匿名データの作成可能性の検討（答申案2 - (1)）

- ・ ペンディングとなっている就業構造基本調査関係の例示については、就業行動に関する研究において各歳別のデータは大変重要であること、具体的かつ重要なニーズが指摘された点については残すべきあること等から、活かす必要がある。
- ・ 就業構造基本調査関係の例示において、「世帯主の年齢を各歳別とする」との記述については、各歳別のデータの重要性は世帯主に限定される訳ではないことから、「15 歳以上の世帯員の年齢を各歳別とする」に修正すべき。
- ・ 全国消費実態調査の年金収入等に関する研究の際にも各歳別データは重要であることから、例示として、就業構造基本調査のほかに全国消費実態調査も追加すべきではないか。
- ・ 全国消費実態調査の例示への追加については、本項目の意見の主眼は複数の匿名データの作成の可能性に関する検討が必要という点にあること、就業構造基本調査は単に一つの例示にすぎないこと、当該作成に当たっては調査客体の特定の可能性等の観点から十分な検討が必要なこと等から、適当でないと考えらる。

ウ 匿名データの作成対象調査の拡大の検討（答申案2 - (2)）

「平成以降に実施したもの」及び「平成より前に実施したもの」という部分については、「平成」のみでは時点が特定されていないので、「平成元年」と修正する必要がある。

エ トップコーディング等が行われた変数の平均値等を提供する措置等の検討（答申案2 - (3)）

- ・ 「回帰分析」という用語は、より一般的な「多変量解析」の方が適当である。
- ・ ペンディングとなっている利用目的を限定した各歳別データの提供に関する記述については、2 - (1) の記述により各歳別のデータの作成の可能性が検討されるのであれば不要である。
- ・ 「メタデータとして利用者に提供する」との記述については、本部会は匿名データの「作成」に関する審議を任務としていることから、「提供」という用語が適当でないため、「メタデータとして整備する」と修正すべき。

「答申（案）」については、上記意見のうち合意された部分について所要の修正を行うことで了承され、3月9日（月）開催の統計委員会に諮ることとされた。

（2）その他

3月9日（月）の統計委員会への「答申（案）」の提示と併せ、今回の全国消費実態調査等に係る匿名データの作成に関する審議の中で出された「公的統計の整備に関する基本的な計画」（以下「基本計画」という。）の推進に参考となる意見として、「訓練用・教育用の簡易な匿名データの作成等」及び「匿名データの作成に係るニーズの把握」について、部会長から報告することにつき、意見交換が行われた。

各委員等の主な意見は次のとおり。

- ・ 訓練用・教育用の簡易な匿名データの作成等に関する検討の中には、当該データの作成のみならず、それを「簡便な手続きで提供すること」も含まれることを明確にした方が良い。
- ・ 訓練用・教育用の簡易な匿名データの作成を直ちに行うことが困難な理由の一つとして、統計法の見直しが必要であることを挙げられているが、同法の運用の中で対応できることもあるので、広く受け取られるよう「法令」の見直しが必要であることとした方が良い。
- ・ 匿名データの作成に係るニーズの把握について、基本計画に関する統計委員会の答申に挙げられている関係学会や経済界等との意見交換の場を利用して行うとのことだが、「経済界」とは一般的に経営者の団体を指すものであり、匿名データの利用者としては考えにくいのではないかと。

統計委員会での部会長からの報告については、上記の意見を踏まえつつ行うこととされた。

総務省政策統括官室から、調査票情報等の利用及び提供に係る制度の概要について、参考説明が行われた。

<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>