

第11回評価分科会 議事録

1 日 時 令和3年6月23日（水）16:00～17:10

2 場 所 Web会議

3 出席者

【委 員】

椿 広計（分科会会長）、岩下 真理（分科会長代理）

【臨時委員】

久我 尚子、山本 渉、美添 泰人

【専門委員】

神林 龍

【審議協力者】

文部科学省総合教育政策局調査企画課係長、農林水産省大臣官房統計部統計企画管理
官統計企画管理官補佐、経済産業省大臣官房調査統計グループ統計企画室参事官補
佐、日本銀行調査統計局経済統計課統計総務グループ長

【事務局（総務省）】

統計委員会担当室：萩野室長、栗原次長、福田補佐、増成補佐

4 議 事

- (1) 令和2年度統計委員会評価分科会審議結果報告書(案)(第8回～第11回審議分)
について
- (2) その他

5 議事録

○椿分科会長 ただ今から第11回評価分科会を開催します。本日も蔓延防止ということで、Webexを使ったネット会議となります。本日用意されている資料につきまして、事務局から簡単に御確認いただければと思います。

○福田総務省統計委員会担当室室長補佐 本日の資料は、令和2年度統計委員会評価分科会審議結果報告書(案)(第8回～第11回審議分)が資料です。参考資料として、第9回評価分科会議事概要及び第10回評価分科会議事概要を準備しております。

○椿分科会長 議事に入ります。本日は、第8回以降これまで議論を行ってきた事項に関する審議結果報告書の取りまとめを行いたいと思います。まず報告書案冒頭から2番の審議結果の4、欠測値の補完に関わる主な方法等についてまでの部分について審議を行います。事務局から説明をお願いいたします。

○栗原総務省統計委員会担当室次長 資料を説明させていただきます。審議結果報告書(案)

の第8回から第11回審議分でございます。おめくりいただいて、3ページに検討の経緯というのがございまして、これが5ページ目まで続いております。この部分は事実関係でありますので、割愛させていただきます。

6ページをおめくりいただきまして、IIの審議結果、1の社会教育調査(民間体育施設)についてです。まず、取組の現状です。社会教育調査のうち、民間体育施設を対象とした調査につきましては、全数で行われているところですが、回収率が約60%で、全部非回答に対して、単一補完やウェイト調整といった欠測値への対応が行われることなく単純合算集計がなされ、結果が過少になっていると見込まれるという指摘がございまして。母集団を明確にした上で、同調査における対応を検討し、結論を得ることが必要であるという課題が示されておりました。

これに対して、文部科学省で研究会を開催し、以下のとおり対応方針の報告をまとめました。

1つ目が、これまでできるだけ広く民間体育施設の全体像を把握する観点から、施設の定義を示してきたが、様々な主体が含まれていて多様過ぎるために、一義的に定義を示すことが困難であり、都道府県・市町村教育委員会において母集団名簿を整える際に差が生じていたことを踏まえて、独自に母集団全体の名簿を整えることは極めて困難なことから、事業所母集団データベースを母集団名簿として活用するというのが1つです。それから、標本調査とすることは適当ではなく、回収数を確保する観点から、引き続き全数調査とするというのが2つ目です。それから、3つ目としまして、都道府県別にウェイトの調整を行っていくということが示されたところです。

これに対して御審議いただいた結果の取りまとめが、2)の評価及び課題解決に向けた今後の取組の方向性のところです。文部科学省が事業所母集団データベースを母集団名簿として活用することとした対応については、母集団が明確化されたことに加え、名簿の更新に関する調査主体の負担軽減にもつながることから妥当であると評価する。

それから、引き続き全数調査としたことや、集計の際の補完方法として、回収率を反映させたウェイトの調整について、全国一律ではなく、都道府県別に調整を行うこととした対応についても妥当であると整理しております。ただし、基幹統計として現状では低い回収率について、上げる努力をすることが必要であると考えられるとしております。

続きまして、2番の建設工事統計調査(施工調査)の関係でございます。建設工事統計調査につきましては、全部非回答を「生産なし」とみなして線形推定が行われていることから、調査結果の過少推計の可能性が指摘されています。

これを受けまして、国土交通省では平成29年度に調査研究を実施しました。その内容としましては、平成24年経済センサス活動調査と平成23年度実績分のこの調査の調査票情報でマッチングを行い、全部非回答の調査票情報を経済センサスとマッチングしたもの、あるいはマッチングしなかったものという形で2つに分けまして、マッチングしなかったものには0値を補完、マッチングしたものについてはウェイトの調整によって補完するという方法を示したということでございます。

これに対しまして、評価分科会では調査研究の成果を高く評価した上で、平成28年経済

センサス活動調査の情報を用いて、調査研究で用いた方法によって補完効果の確認を行う必要があると指摘したものでございます。

この評価分科会による指摘を踏まえ、国土交通省で改めて令和元年度及び2年度に調査研究を実施しました。この研究では、平成28年経済センサス活動調査と、平成27年度実績分の同調査の調査票情報でマッチングを行いました。それに加えて、行政記録情報の経営事項審査結果との照合を行い、全部無回答の調査票情報を、i)として、経営事項審査結果との照合により補完できる調査票、それからii)として、経済センサスとマッチングした調査票、それからiii)として、そのいずれともマッチングしない調査票の3つに分別しました。

そして、経営事項結果とマッチしたものは、行政記録情報のデータを用いて補完、いずれともマッチしなかったものは0値を補完、経済センサスとマッチしたものはウェイトを調整する方法によって補完を行うという方向性を示したところです。この方法で補完したことにより、経済センサスの結果と比較してもほぼ合致することが確認できました。

これに対する2)の評価でございませう。今後の対応としまして、行政記録情報を活用して補完を行った後に、残った部分に関しまして、経済センサスとの照合結果を踏まえたウェイトの調整による補完を行うこととしていることについては、妥当であると評価する。なお、この手法については、行政記録情報を活用した補完方法ということで、他の統計にも参考となる事例ではないかと考えられると整理しております。

続きまして、3番目の経済産業省生産動態統計調査についてです。生産動態統計調査については、欠測値が発生した場合には、過去の回答結果を使用した単一補完(LOCF)が明確な使用期限を定めることなく行われているということがございました。回収率は9割以上を保持しているのですが、総務省の支援も得ながら計画的なシミュレーションによってデータの使用期限も含めた補完方法の検証を行うことが必要であるということが課題として挙げられていたものです。

このため総務省の統計研究研修所の協力を得て、同調査の自動車等の調査票情報を基に、クロスバリデーションに類似したシミュレーションが行われたものです。そのシミュレーション結果を受けて、第3回評価分科会では、LOCFは平均値を補完する方法や、年の伸び率で補正して補完する方法、前年同月比で補正して補完する方法など、他の方法と比較しても補完効果はほぼ変わらないことが報告されるとともに、シミュレーション結果を踏まえて、総務省から提案として示された同調査におけるLOCF法の利用に関する考え方について審議が行われたものです。

その考え方は3点ございませう。1番として、長期的にLOCFを行えば、精度悪化のリスクが高まるものの、数年程度であれば極端に悪化するものではないことから、自動車等に含まれる品目のうちで、金額の変動が大きい品目であって、補完期間が5年に至るものにつきまして、5年ごとに実施される経済センサス活動調査を活用して、当該事業所の伸び率等でLOCF値を補正するなどの対応が1つです。

それから、自動車等に含まれる品目のうち、複数年にわたって金額の大きな変動が続いているような品目については、2年以内に同じ品目内の伸び率等を乗じることで、LOC

F 値を補正いたしますとともに、5年ごとの経済センサス活動調査の当該事業所の結果を用いて補正するなどの対応を行うというのが2つ目です。

3つ目が、検証を行っていない自動車等に含まれる品目以外の品目につきましても、上記のルールを暫定的に導入してはどうかというものでございました。

これにつきまして、評価分科会での御審議では、シミュレーション結果に関しまして、これまでの結果で判断しますと、自動車等の品目に関しては、L O C F 法は一定程度有効に機能している、上記で示したような対応は妥当であると評価したところでございます。

ただし、多数の品目で検証を行うまでの必要はないものの、自動車等の品目以外の二、三の品目において同様の検証を行って、その結果を踏まえて判断することが適当であるという指摘がなされたところです。

この指摘を踏まえ、改めて総務省統計研究研修所の協力を得て、電子部品及び段ボールを取り上げ、同様のシミュレーションが行われました。このシミュレーションの結果について、経済産業省ではその評価として、時間が経過すれば乖離は拡大するものの、一部の特殊な品目を除き長期的に乖離が極端に拡大する傾向は観測されなかった、数年程度であれば一部の特殊な品目を除き L O C F 法は有効に機能すると評価した上で、今後報告をいただけない事業所への依頼や要請を最大限実施することが最も重要であると認識を示し、報告をいただけない場合は L O C F によって欠測値補完を行うとしたところです。

変動の大きな品目への対応としては毎年の年間補正の中で行い、安定した推移をしている品目への対応としては、経済センサス活動調査における、生産動態統計の無回答事業者と同一の事業者規模等の層に属すると思われる事業所集団の傾向を用いて、L O C F 法による値を5年ごとに見直すという方向が示されました。

これに対する評価ですが、今回のシミュレーション結果を踏まえ、数年程度であれば、一部の特殊な品目を除いて L O C F 法は有効に機能していると評価した上で、今後の対応として、変動の大きな品目への逐次対応は年間補正の中で行い、安定した推移をしている品目については経済センサス活動調査における、無回答事業所と同一の層に属すると思われる事業所集団の傾向を用いて、L O C F 法による値を5年ごとに見直すとしたことは妥当であると評価する。今回のシミュレーションを通じて、時系列的なデータに対する補完手法を比較検討する中で、L O C F 法を利用することの効果が確認されたことを評価するという形で、分科会での御審議を整理いたしました。

続いて4番の欠測値の補完に係る主な方法等です。こちらは第9回の分科会で御審議いただいた事項です。公的統計を作成するための統計調査において、調査の実施段階においては、調査客体の確認や督促なども含めて、調査票への記入及び回収の確保に努めているところですが、一部の調査項目が未回答である場合や、回答そのものが得られずに、全ての調査項目が欠測となった場合には、結果の有用性を確保するため、統計的な補完を適切に行うなどの対応が考えられる。

一方で、欠測値については、統計調査ごとに欠測の発生状況や補完に利用できるデータなどに違いがあるため、特定の補完方法の適用など一律な対応は困難な面がある。このため、統計調査ごとの状況を踏まえた適切な対応が重要となりますが、補完を行うに当たっ

ての主な方法、手順や、利用上の注意点など、実務上参考となる事項を整理してまとめて示すことは、公的統計の正確性や有用性の確保を促進する上で有意義と考えられます。

このような観点から、これまでの分科会における審議や各府省における取組状況等も踏まえて、事務局において作成した資料を基に評価分科会において審議を行い、後ろに付けております資料5のとおり取りまとめたと記載しております。

次のページです。統計調査において、欠測値をできる限り発生させないようにすることが最も重要であるが、最終的に発生した欠測値への対応を行う上で、今回の取りまとめが各府省の実務担当者に広く共有され、欠測値への補完等の方法について理解を深めるための参考となるとともに、補完等の対応が適切に行われ、かつその適切性が国民に対して十分な説明がなされていることで、公的統計の正確性や有用性の一層の確保に寄与することを期待するという形でまとめさせていただいています。

それで、通し番号49ページ、資料5を御覧いただければと思います。こちらは御審議いただいた御意見について、基本的に反映させていただいたところですが、修正のポイントだけ簡単に説明させていただきます。

まず、目次でございます。こちらは全体像がより分かりやすくなるよう、括弧数字のレベルまで記入いたしております。構成として、2番の主な手順、3番の主な方法ということで、原案のときから順番を入れ替え、手順を先に持ってきています。

それから、文章の中で用いている用語の関係です。まず、客体と標本というところでは、こちらの整理としましては、文脈によりまして、サンプルとしての集合体という意味の場合には標本としています。それから一つ一つのレコードの意味の場合には、客体という形で改めて整理させていただいております。

それから、項目と変数という御指摘もいただいたところでは、こちらは調査票上の調査項目という意味で使われる場合には項目、それから推計上用いる変数といった場合には補助変数とするという形で再整理させていただいております。

それから、平均値という言葉と平均という言葉の使い分けについてです。母集団の平均という意味では平均とし、補完として代入に用いる値という意味では平均値と整理しています。

それから、本文で一部追記等をした箇所です。補完を行う前に欠測が発生しないようにすることが必要だという趣旨がより分かるようにするという御指摘を受けまして、3ページの(2)、1つ目の「○」として、そこは単独で立てているところです。

同様に、「はじめに」の冒頭でも、回収の確保に努めた上で欠測が発生した場合にはという形で再整理をしています。

それから、3ページ目の下段で、マイクロデータ利用面での意義についても追記しています。

以上のような点について、御指摘を踏まえ修正を行わせていただいたところでは、事務局からの説明は以上でございます。

○久我臨時委員 2点ありまして、まず、1番目の社会教育調査で、6ページの一番下に、低い回収率について上げる努力をしてくださいという記載があり、これはいいと思います

が、続いて7ページの建設工事統計調査も基幹統計ですが、こちらも回収率が同様に6割と低いです。こちらも方向性のところで、記載を同様に合わせた方がよいのではないかと。そもそも低い回収率を上げる努力をしましょうというところです。

あともう一点、以前にもこちらの会で未回答と非回答と無回答という表現が混在しているという話をいたしました。そのあたりも可能な範囲で統一した方がよいのではないかと。例えば、資料10ページの欠測値の補完に係る主な方法等についてというところの未回答という表現、資料編の中でも未回答という表現をされていて同じ話題の中で統一されていますが、その2点が気になりました。

○栗原総務省統計委員会担当室次長 基本的に文書は評価分科会の際の議論を踏まえて整理していることもありまして、書きぶりが若干違っている部分はありますが、御指摘いただいた点を踏まえ、もう少し整理を検討したいと思います。

○樫分科会長 表現ぶりは、調整し、御指摘を反映させていただくという形で良いのではないかと。非回答、無回答のことはどうですか。

○栗原総務省統計委員会担当室次長 そちらの用語も御指摘を踏まえて再度整理させていただきます。非回答につきましては、平成30年の精度検査報告書の中で、そのように書かれていたことがありまして、こういう表現がされているということです。それも踏まえた上で、今回の文章の中での表現ぶりは統一するかどうかも含めて、座長とも相談させていただきたいと思います。

○樫分科会長 分かりました。いずれにせよ修正の必要性も含めて、御指摘の点を踏まえて、事務局と私の方で一任させていただければ幸いです。

○美添臨時委員 本日の報告で特に4番の「欠測値補完に関する主な方法等について」とその附属資料はとてもよくできていると思います。私がこの分野を最初に勉強したのは、1970年代にアメリカのセンサス局と私のハーバードの先生方が一緒にまとめた研究書でした。その後も欠測値補完の方法については各国の統計局が研究成果を提示しています。大学の世界だとこの分野を専門にする人はあまり多くなかったので、今回の資料はすばらしいものができましたから、英語版も出したら良いのではないかと。

日本の統計に対する批判はよく聞きますが、日本の統計はとてもよくできていると考えています。統計局は遠慮し過ぎて、海外に対する情報発信が少ないと思います。

今日は日銀もいらしていると思いますが、日本銀行はきちんと出しています。それと比較すると、統計局はせっかくここまで作業して報告書を取りまとめたのだから、是非、海外に対する情報発信も考えていただきたいというのが1つのお願いです。

あと細かいことで、平均と平均値とか、無回答と非回答、欠測と欠損値という表記が出ています。いろいろな表記があるので、統一したものを事務局で考えていただければ良いのですが、例えば、英語にすると、平均と平均値は両方とも mean ですね。私の先生の P.フーパーのように、アベレージという表現を好む数学者もいないわけではないけれども、普通は mean で、日本語で平均でも平均値でも私は同じ意味だと思っています。

少し心配なのは、文部科学省の学習指導要領では平均値と書くこととしたが、学会では平均も広く使われているため、細かな違いを気にする人もいるらしい。このような事情は

統計局がよく御存じだと思いますので、誤解がないような整理を、よろしくをお願いします。

非回答と無回答は、これも英語でいえばどちらもノンレスポンスです。

欠測値は、欠測の方が良いと思います。欠損という人もたまにいますが、それは会計用語だからいけないと聞いたことがあります。少なくとも統計用語のミッシングオブザベーションは観測の「測」なので、欠測の方が欠損よりいいと、学会では何回かそういう提案をしたことがあります。これについても、英語は同じですから、無用な混乱さえ起きなければいいと思います。

それと、文部科学省、国土交通省、経済産業省と3つの事例について検討していただいた。それぞれすばらしい書きぶりで、対外的にはこれでいいと思います。実質はどうかというと、まず、国土交通省は令和元年と2年にかかなり時間をかけて調査研究を実施しています。報告を伺った中でも、欠測値の補完方法について、高い能力を持った担当者がいました。国土交通省に高度な統計の知識を持った職員がいて、このような素晴らしい分析をしたことに、本当に感心しました。このような職員が育っているなら、回収率が低いということに対して、それなりの対応もできる、今後も安心してお任せできるという、率直な感想です。

それから、経済産業省の生産動態は、総務省の全面的な協力を得たということですが、自動車等を分析した後、更に段ボールと電子部品に対して本当に丁寧な分析を実施していただきました。今後も安心してこの方法でやっていただければ良いと思います。

3つ目の文部科学省は、少し本業が忙しいのではという気がします。方向として書いてあることはそのとおりで、回収率の向上は文部科学省が努力すればある程度は解決する問題だと思います。一方で、大事な問題は母集団の整備をどうするのかということです。母集団の整備と利用方法については、文部科学省に対してのお願いですが、総務省の知恵を借りるか、若手の研究者に声をかけて、内部の研究会を設置していただきたい。学会の、特に若手の知恵を借りるのが得策だと思います。職員だけで検討するのは負担が大きいと思います。

最近この評価分科会で出た国税庁の調査では、報告当時には研究会は開催されていなかったのですが、評価分科会の後で、若手も入れた充実した検討の機会を作っています。そこで、短期間に課題が整理できて、大きな改善が実現しつつあります。

文部科学省もそうやって外の方の知恵を借りることを是非試みていただきたい。内部だけだと負担がかかると思います。統計の専門家の知恵はそれほど費用をかけなくても提供してもらえます。アメリカ統計学会では、統計学者の無償奉仕は統計の評価に結びつかないからコンサルティングなどでは正当な対価を取れと言うのですが、日本ではまだそういう雰囲気ではなくて、統計の改善のためには貢献してくれる研究者がたくさんいるわけです。結果としていい統計ができれば、研究にも使えるわけですから。財務省の法人企業統計などはそういう例ですが、文部科学省の統計も貴重な統計だと思うのです。その品質の維持・向上のために外部の知恵を借りる工夫をしていただけると、文部科学省内部としても、人材に過重な負担をかけずに改善を図ることができると思っています。これは是非ご検討をお願いします。

○樫分科会長 報告書の背後にまだ問題が存在している。これだけのものができたので、海外発信を行うことに関しては、おっしゃるとおりのようなことができればと思います、英訳に関しては、マンパワーとカリソースの問題を別途考えなければいけないのではと思います。

○栗原総務省統計委員会担当室次長 御指摘も踏まえまして、情報発信の充実という面もよく考えて対応してまいりたいと考えております。

○樫分科会長 これだけ立派な報告書が出てきたら、研究者側でボランティアとは言いませんが、何らかの研究班がこういうものをきちっと海外発信するために支援する体制があってもいいと思います。

あとは、府省の統計調査の欠測なり、情報設計ももちろんそうなのですが、そういうものをアカデミアが支援できるような研究会組織の組織体がいろいろな府省に対してあるようにする。特に文部科学省に対して御指摘がありました。これも今日のメインの議題とは違うかもしれませんが、御指摘があったということは記録に残しておいていただきたい。

○栗原総務省統計委員会担当室次長 御指摘いただいた点は文部科学省さんにもお伝えするようにしたいと思います。

○樫分科会長 それでは、先ほど御指摘がありましたように、文言の標準化というか、統一ができていのかどうかに関しては、事務局と私でもう少し検討させていただいて、報告書自体の流れといたしますか、様式は今日いただいたものを承認するとして、その種のエディティングの作業は、もうしばらく行わせていただきたいと思います。その作業自体は、申し訳ありませんけれども、事務局と私で相談させていただくということによろしいでしょうか。

用語の標準化というのはなかなか大変なものですけれども、ISOの統計用語とか学術用語とか20年ぐらい携わりましたが、英語で考えたら良いのですが、それをまた翻訳するとなると日本の中で統一されていないという話があったりして、非常に苦労するところが多い。微妙な違いがあるために、英語では同じものが日本語では違っていて誤解を呼ぶというのは残念ですから、言葉の標準自体は作って行って、お互いに言葉が通じるようにするという事はやりたいと思います。

○美添臨時委員 もう一言、前から気になっているのが、サンプルサイズの日本語訳です。統計局が昔編集した事典には、標本の大きさが正しく、標本数は間違いだと書いてあり、学会の事典にもそう書いてありますが、統計局のホームページには、サンプルサイズのことを標本数と書いてある例が少なからずあります。調べましたら文部科学省もそうです。

統計学者の中で、標本数という表現は間違いだと常々言っている人が、あるとき標本数という用語を用いたので、この人も標本数と書くのかと思って、たずねたことがあります。そんなことは言っていないと反論されたことがあります。このように、本人はうっかりして使ってしまうぐらい世の中で使われている言葉なのです。標本数の英語はサンプルサイズです。two-sample testと言う場合、標本が2つある、それぞれサイズがいくつという。ここは学会と統計局で相談させていただいて、標本数には2つの意味があることにするなどの整理をしていただけないでしょうか。サンプルサイズの日本語訳で「標本の大き

さ」という表現が使いにくいということだと思います。片仮名でサンプルサイズと言うなら、それはそれで良いのですが、本や原稿を書くたびに悩みます。評価分科会で提案するのがふさわしいのかどうかよく分かりませんが、学会と統計局、政府統計とすり合わせていただけないか。

○神林専門委員 その点については、経済学でもかなり混乱してしまっていて、事実上、標本数という言葉と、観測数、オブザベーションの数というのでサンプルサイズを対比させて、標本数のときは実験の数という言い方をして、エクスペリメントの数とか、エクスペリメントの種類とか、そういう言い方をして、用語を使い分けているのですが、標本数という言葉を使えるようになったらもっと楽になるだろうと思っています。そこははっきりさせた方がいいというのは、確かにそのとおりだと思います。

○美添臨時委員 統計学会と経済学会と総務省を入れて相談すると良いですね。

○神林専門委員 ええ、合わせて何か声明でも出す、ノートみたいなものを出して、これでいきましょうといったもの。

○美添臨時委員 これは本当に基本的な用語ですから。

○神林専門委員 実は学生が全く混乱していて、理解していないことが多いです、サンプルとオブザベーション。これはこの一番で、多分1年生ぐらいで勉強すると思いますが、実際の論文や論考を読んだりしていると、そちらが混乱しているので、昔勉強したことが、あれ、そうだったか、どっちだろうとどんどん混乱していってしまう。自分が見ている範囲ではそれがもうかなり進行しています。

○椿分科会長 標本数は非常に深刻な問題で、さっき申し上げましたように、サンプルサイズに関しては、日本工業規格というか、産業界では標本の大きさというものを使わなければいけないことになっていますけれども、実態としては、例えば、企業の中では、標本数とかN数とかが圧倒的に多いということは承知しています。品質マネジメントシステム監査に入るときには、JIS規格とかISOに準拠しないと怒られてしまうので、正しい用語を使っていますけれども。確率統計用語の日本工業規格を作るときは、学術会議とか統計局に御意向を確認することにはなっているのです。

○美添臨時委員 皆さんの知恵が必要ですね。もう一つ、データセットという言い方もあるのです。データセットというのは、サンプルとほとんど同じ意味、エクスペリメントと同じ意味です。オブザベーションと違っていいのかというと、実験した場合はオブザベーションではないのです。あれはエクスペリメントの対象数なのです。ですから、本当に難しいです。英語は何でこんなに簡単なのだろうと思います。日本語が定着してしまったのと、とにかくサンプルサイズの日本語訳で苦労したのが問題ですね。そろそろ何とかしてほしい。

○椿分科会長 さっき言いましたように、産業界は国際標準で統一しています。

○美添臨時委員 これは学会も巻き込んでお願いします。ちなみに、調査業界、JMRA—日本マーケティング・リサーチ協会では、客体のことをサンプルというのです。あなたはサンプルに選ばれましたと。1人ずつがサンプルなのです。それもその業界用語だと思えばいいのですが、主要な民間統計のメーカーですから、公的統計、民間統計、学会、こ

の辺で何か相談していただけると、安心して使えるなど前から思っています。

本来は誰の業務かはわかりませんが、総務省としても、公的統計の用語ですから、心にとめておいて、結論を出していただいたら、従います。よろしくお願いします。

○椿分科会長 もともと学術会議の統計数学用語が終戦直後にできて、それに併せてJIS規格ができて、当時はそれなりの規律があったのです。標本の大きさをここで議論するとは思わなかったのですけれども、実に困った問題だと思っています。

これを評価分科会でやるかどうかということは別として、さっき言ったように、用語標準は誤解を呼ぶ部分があって、サンプルはもちろん、サンプル、標本という訳と、もう一つは重要なものとして、試料という概念があります。その辺も実は分野によって非常にばらついています。国内だけでなく海外でも実はばらついているものがありますけれども、そういうものをうまく整合させていただければと強く思います。要は、誤解がないようにしていただきたい、分野によって誤解がなくなるようにしていただきたいということが強くあります。これは非常に重要な話だと思えますけれども、概念に定義を与えて、用語をきちっと定義する手続きです。非常に貴重な提言で、重い提言であることも事実だと。教育用語と産業界用語と統計自体の用語ということでもうまくすり合わせができればと思います。

ここまでの点でほかにコメントはありますか、よろしいでしょうか。

それでは、最終的には one more effort で、言葉の統一を行っていくということは、私どもの作業としてやらせていただいた上で、一応先生方に確認を取って、こういう形に統一しましたということをご報告させていただければと思います。

それでは続きまして2の5、令和3年度における評価分科会の取組について、審議に入りたいと思います。これも事務局から説明をよろしくお願いします。

○栗原総務省統計委員会担当室次長 資料 11 ページの5番、令和3年度における分科会の取組についてということで、2点挙げさせていただいております。

1つ目が、報告書のフォローアップでございまして、欠測値の審議でまだフォローアップが行われていない経済産業省企業活動基本調査につきまして、その進捗、取組状況を聴取した上で審議を行うということで、1つ目に挙げさせていただいております。

それから2つ目でございます。母集団情報の利用状況についてということで挙げさせていただいております。こちらはこれまで欠測値に対する対応といたしまして、未記入、未回収などへの推計上の対応という面から御審議いただいているところです。その中で、先ほども出ましたとおり、基幹統計調査でも母集団情報の利用において課題等があるものも一部見受けられたところです。正確な統計の作成という観点からは、調査のベースとなります抽出のフレームとして適切な母集団情報が利用されていることも重要だと考えられるところです。

また、調査対象の抽出のための母集団情報の適切な利用という点については、各調査にも共通する重要な課題でもあるということで、椿座長とも御相談の上で、令和3年度のテーマの一つとしてはいかがかということで、記載しているものです。

○神林専門委員 (2)の「母集団情報の利用状況について」というのは、具体的にはど

のようなことを念頭に置いていますでしょうか。

○栗原総務省統計委員会担当室次長 基幹統計の中で、具体的に母集団情報、抽出フレームとして何を使っているかを、まず事務局でよく事実確認をした上で、事業所系であれば事業所母集団DBもありますので、その利用可能性なども含めて必要な御審議をいただくことをイメージしています。

○神林専門委員 ということは、抽出プロセスも含んで検討するということですね。

○栗原総務省統計委員会担当室次長 標本の抽出方法までは想定しておりません。フレームとしての母集団として何を使っているか。例えば、事業所系の調査であれば事業所母集団DBを原則として使いなさいということになってはいますが、実態としてどうなのか、もしきちんとした理由があって事業所母集団DB以外を使っているのであれば、それは一つの選択です。先ほどの社会教育調査もそうですが、精度を高める上で事業所DBを使っていくという選択の余地があれば、そうした方向で少し考えていくとか、特にフレームの利用に主に重点を置いて、このように書かせていただいているところです。

○神林専門委員 分かりました。

○美添臨時委員 神林先生の理解は、標本の抽出枠まで含めてですかという質問でした。利用状況と書いてあるのだから、事務局の説明のように、母集団情報は何かから作成しているか、メンテナンスに関する状況プラス利用とすれば、募集団枠だと思ったのですが、そうではなくて、メンテナンスの話だけだということのだったら、利用状況という言葉は誤解を招きます。利用と言われたら、私も標本の抽出枠だと思ってしまいます。

○樫分科会長 何を母集団としているかということがメインです。

○美添臨時委員 サンプルフレームとして使うのは当たり前だから、母集団について確認する、そういうことですか。

○栗原総務省統計委員会担当室次長 少し補足させていただきます。先ほどの社会教育調査の場合は、これまで教育委員会が管理する名簿で行われていた。ただ、カバレッジとか正確性とか、そういう面に少し課題があって、今回事業所DBを用いる方向に変えていくことになっています。

同じような観点から、事業所DBを使っていない調査もあると思うのですが、行政記録など母集団情報がきちんとしたものが使われているのかどうか、データの更新や改廃の状況とか、もろもろ正確な調査を行う上で必要な情報としてのフレームという観点を考えております。そういう意味では、美添先生のおっしゃる管理の観点も含んでいるものとも言えるかと思えます。

○美添臨時委員 言っていることの意味は分かりました。そうであれば、表現を変えていただいた方がいいと思います。今おっしゃったことは、これまで統計審議会あたりで言っていた表現を使いますと、母集団名簿の整備状況です。私の記憶では、これをずっと使い続けてきたと思います。母集団名簿の整備状況といえば、今の内容が反映されると思います。それだと何かおかしいですか。

○樫分科会長 基本的に、例えば、事業所母集団のデータベースを使っていたのがメインの方針だと思うのですが、それ以外のフレームを歴史的に使っている調査が多々ある

わけです。

○美添臨時委員 事業所企業だったら、経済センサスの使用に移行しつつありますが、それも確定ではないし、経済産業省、通産省は昔から工業統計名簿とか、商業統計名簿を別にメンテナンスしてきたわけですし、国税だって別な名簿はあるわけで、経済センサスだけではないですね。しかも、世帯、個人に関しては国調絡み、いろいろなものがあるわけですから、利用状況という書き方は、私はちょっと理解できません。

○樫分科会長 経済産業省はもう既に先程申し上げた方向に動いています。

○神林専門委員 労働系だと雇用保険の事業所の名簿が使われますね。厚生系だと、病院の名簿があって、必ずしも事業所企業の母集団データベースが使われていないというのは事実だとは思いますが、どちらがよいかということに関しては結論が出てないと思いますが、結論は出ているのですか。

○樫分科会長 統計委員会の基本計画ではどうなのですか、確認ですが。

○神林専門委員 母集団データベースに統一していこうということになっていますか。

○栗原総務省統計委員会担当室次長 事業所系について言いますと、事業所母集団データベースを調査の特性に応じながら、原則として使っていこうという方向が出されています。したがって、調査ごとにいろいろな種類の名簿が使われている中で、そうした方向性を踏まえて、今後全体としてどのように対応していけばよいのかということがあるということでございます。

○神林専門委員 そうしたら、各府省庁が今使っている独自の母集団名簿と事業所母集団データベースとの違いをはっきりさせた方がいいですよ。

○樫分科会長 そこはおっしゃるとおりだと思います。

○美添臨時委員 事業所企業に関してはデータベースを充実したいということで、それは方針として分かります。だから、事業所企業に関しては、経済センサスの利用状況を知りたい。そこは分かりました。しかし、個人とか世帯はどうするのかまた別な話がありますから、この書きぶりだと誤解を招くと思います。何か表現を考えていただけませんか。

○神林専門委員 世帯系のデータ統計は対象ではないですね。基幹統計調査とどんと書いてありますけれども、例えば、労働力調査なんかは関係ないですね。関係あるのですか。

○美添臨時委員 関係あります。重要な基幹統計ですよ。

○神林専門委員 母集団データベースとは関係ないですね。

○美添臨時委員 母集団データベースの事業所から抽出しろと言われてたら、できないことはない、まあ、労働力調査は世帯を対象に調べますけれども。

○山本臨時委員 山本です。母集団情報という言葉から、全数調査の統計も全て扱うのかと思ってしまいました、情報というのは結構曖昧な表現なので。もう一つ、先生方が今御議論されていた母集団データベースというのも、統計局のホームページを検索してみましたら、正確には、事業所母集団データベースです。(2)は、「事業所対象の単位の調査に関して」と付記し、「事業所母集団データベースの利用状況及び他の母集団名簿を使っている調査との比較」といった項目になるような気がします。いずれにせよ統計調査における母集団情報という用語が私は理解できなくて、先生方の話を聞いて勉強していました。

もう少し書き加えていただいた方がよい気がします。

○樫分科会長 基本的にフレームを言っていると思います。

○栗原総務省統計委員会担当室次長 事業所系統計調査であれば、事業所母集団DBがありますので、それとの関係になるのですが、事務局の方でまず全体を点検した上で、課題を洗い出すということで、ここでは特に事業所だけに絞らない書き方にさせていただいております。

○美添臨時委員 「基幹統計調査の標本抽出における母集団情報の利用状況」と書いてあります。余計混乱してしまう。山本委員の整理では、基幹統計調査、「標本抽出における」を取っ払って、「母集団情報」のところに、「事業所企業母集団データベースの利用状況」、それでしたらびったりですね。しかし、世帯調査は入らないが、それで良いのですか。

○山本臨時委員 そうすると、利用状況ですので、使っていないのであれば、どういう名簿を使っているかと対比ができるので、使っていなくても良いと思うのですよ。

○美添臨時委員 事業所母集団データベースの利用状況ですよ。

○山本臨時委員 使っていないときに、使っていませんではなくて、それだったら、どんな名簿を使っているかということが分かると良いと思います。

○美添臨時委員 それは明確だけれども、そうすると、標本抽出は関係ないし、世帯と個人も関係なくなってしまう。それで良いのであれば、明確にはなると私は思います。

○山本臨時委員 情報というのだけが曖昧だったので、「事業所母集団データベースと母集団名簿の利用状況について」とか、「等」を入れて、「母集団名簿の定義及び活用状況について」のような感じでもよいような気はします。

○樫分科会長 その調査が必要だということでしょうかね。

○栗原総務省統計委員会担当室次長 そうですね。ただ、例示を事業所母集団等としてしまいますと、事業所系だけを対象にしているように見えてしまうような気がします。

○山本臨時委員 「等」はなくてもよいと思います。先ほど美添先生が気にされていたように、情報という言葉が何を指していて、「利用状況について」というのが、何をすることなのかもう少し具体的だと誤解がないのではないかと感じた次第です。

○樫分科会長 「母集団情報」という言葉を使う必要はほとんどなくて、「母集団名簿」で全く問題ないですね。

○神林専門委員 確認ですが、世帯系の統計も入るということですね。

○樫分科会長 もちろん入ります。世帯系、事業所系ということに限らず、母集団名簿は何を使っているかということ、特別な母集団名簿を使っている場合には、その理由、妥当性とか、それに対する管理はどうなっているかということを知りたいということです。

○美添臨時委員 もう一つ、「標本抽出における」は残すのですよね。さきほどの栗原次長の説明だと、標本抽出はあまり中心ではないように聞こえたのですが、これは残していいですか。

○樫分科会長 「標本抽出」は不要だと思います。基本的には母集団だと思います。

○美添臨時委員 「標本抽出における」があれば、サンプリングまで考えなければいけませんね。神林先生の提案というか、さっきの確認がそうでしたね。

○樫分科会長 今回の問題意識は母集団名簿だけですよね。

○栗原総務省統計委員会担当室次長 分かりました。標本抽出というのと、抽出の仕方までがターゲットのように見えるというのは、確かにそのとおりかもしれませんが、「標本抽出における」というのは、とらせていただく方向で考えたいと思います。

○樫分科会長 むしろ独自の母集団名簿を使っていることの妥当性とその管理がきちっとできているかどうかということについて関心がある。そういう形の文言にさせていただくこととします。非常に重要な点だと思しますので、もう一度見ていただくという操作をしたいと思います。よろしいでしょうか。

非常に貴重な意見をいろいろ頂戴したと思います。(2)の「統計調査における」については、「母集団情報の利用状況について」とのタイトルも今のような趣旨で修正した上で、「標本抽出」というところも削除して、あくまで基幹統計調査の母集団名簿がどうなっているか、標準的なものになっているのかどうか、いわゆる一般的な管理がされているものになっているかどうかということと、独自の母集団名簿を使用している場合にはその管理状況がどうなっているかを検討するというのを、令和3年度における評価分科会の取組にするということで、今のような修文を行った上で、もう一度先生方に見ていただいて、了承いただくという形にしたいと思います。

修文に関しては、今のような形にしたいのですが、何か御意見があれば、もう一度確認しておきたいと思います。サンプリング調査に限らず、完全に母集団の話と管理の話にと私自身は認識していたので、確かに文言的には先生方の御指摘のとおりだったと思います。よろしいでしょうか。

それではそのような修正の上で、一応この部分を確定させていくようにしたいと思いますので、よろしく願いいたします。文面につきましては、事務局と私に一任いただいた上で、もう一度先生方に見ていただくという操作をしたいと思います。どうぞよろしくお願いいたします。

平成30年9月に統計委員会で決定した評価分科会の審議事項に関する委員会の議決によって、評価分科会の所掌事務に関する事項については、この評価分科会の議決をもって統計委員会の議決、そういう原則があるとのことでした。したがって、報告書としても、統計委員会のホームページに掲載、公表という手順を取ることになります。

先ほど言いましたように、今日いただいた意見で若干修文することについては、作業は私どもで行って、先生方にもう一度確認いただければと考えております。その作業方針に関して一応御了承いただいたという整理にしたいのですが、よろしいでしょうか。

どうもありがとうございます。先生方、この際何か御意見があれば、岩下先生は2日間連続で、御意見を伺うチャンスもなかったもので、何かあれば一言よろしく願います。

○岩下分科会長代理 いろいろな分科会に出ていて思うのですが、何か統一したルールがしっかりしていないというのがいろいろなところで起きていて、用語もそうですし、英訳、和訳の話も何度もいろいろなところに出ていて、それは誰がどうしたらうまくまとめられるのかということが、統一した共通の課題かと思いました。

○樫分科会長 英訳の問題はほかのところでも先生と一緒に聞いたような気がします。

基本的に評価分科会のミッションがどこまでか、例えば、さっきの国際発信みたいなのは、日本の統計の中でどこが担うのかとか、実はその辺が仕組みとしてはまだ確立していないのです。先生方に今日は非常に貴重な意見いただきました。報告書に反映させた上で、先ほどのような公表、公開に向けていきたいと思います。

それでは、本日の分科会は、ここまでとさせていただければと思います。最後に、事務局から連絡事項をよろしく願いいたします。

○福田総務省統計委員会担当室室長補佐 次回の日程につきましては、後日また日程調整をさせていただきたいと思っておりますので、よろしく願いいたします。

○樺分科会長 それでは、本日はこれまでといたしたいと思います。今日はどうもありがとうございました。