

## 第13回評価分科会 議事録

1 日 時 令和4年5月17日（火）10:00～11:30

2 場 所 Web会議

3 出席者

【委 員】

菅 幹雄（分科会会長）、川崎 茂（分科会長代理）

【臨時委員】

神林 龍、久我 尚子、山本 渉

【説明者】

国土交通省：総合政策局情報政策課交通経済統計調査室長

【事務局（総務省）】

統計委員会担当室：萩野室長、小山次長、栗原次長、福田補佐、増成補佐

4 議 事

自動車輸送統計調査（全数バス調査）における欠測値補完手法について

5 議事録

○菅分科会長 ただ今から第13回評価分科会を開催いたします。本日は、土屋専門員が御欠席です。

本日は、自動車輸送統計調査の全数バス調査における欠測値補完手法についての審議となります。

同調査の欠測値補完手法については、第5回評価分科会（令和2年1月27日）において一度審議が行われているということです。その後、諸般の事情から、再検証が必要となりましたので、今回その結果を国土交通省から御報告いただくというものです。再検証が必要となりました理由も含めて、後ほど国土交通省から御説明いただきます。

それでは、本日用意されている資料について、事務局から簡単に御確認をお願いします。

○福田総務省統計委員会担当室室長補佐 資料といたしましては「自動車輸送統計調査（全数バス調査）における欠測値補完手法について」が資料となっております。

○菅分科会長 それでは、議事に入ります。

自動車輸送統計調査の全数バス調査における欠測値補完手法について、国土交通省から説明をお願いします。今回、再検証が必要となりました経緯も含めて、よろしくお願いたします。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 本日、御説明のお時間をいただきまして、誠にありがとうございます。

資料の説明の前に、今回、改めて説明に至った経緯をまず最初に御説明をさせていただきまして、資料の説明に移らせていただきたいと思います。資料の2ページを併せて御覧いただきながら、お聞きいただければと思います。

まず、サービス統計・企業統計部会と評価分科会の説明に違いが生じてしまった経緯ということでございます。自動車輸送統計調査につきましては、第3期の基本計画におきまして、調査手法等に関しまして見直しの必要性を指摘されておりました。このことから、令和2年度から新たな調査手法による実施に向けまして、検討を行ってきたところです。

そのような中で、全数バス調査、これは事業所の全数調査ですが、全数バス調査の欠測値補完手法については、平成28年度の統計法施行状況に関する審議結果報告書の統計精度検査関連分におきまして、過去の回答結果を使用した単一補完を実施する場合の使用データの期限のないものというものに整理がなされまして、第3期基本計画において、自動車輸送統計調査そのものの見直しが指摘されていることから、当該見直しと整合するよう、並行してデータの使用期限も含めて検証を行うことが必要であるというふうに指摘を受けたところです。

自動車輸送統計の見直しに係る承認手続を行った際、全数バス調査の見直しと併せて、営業用トラック調査の見直しが行われまして、その推計において採用することとしていた手法と整合が図られるよう、全数バス調査においても層内同一事業所の合計値を前月との伸び率で推計する手法に変更することとし、この内容をサービス統計・企業統計部会において説明をさせていただきました。

また、その後、欠測値補完に関する審議が評価分科会において行われることとなりまして、その中で全数バス調査の欠測値補完手法について、審議結果報告書における指摘を踏まえまして、当省で実施しました欠測値補完に係る調査研究業務の分析結果に沿って説明をさせていただいたところです。その結果、サービス統計・企業統計部会と評価分科会での説明に違いが生じたところがございます。当時、管理職や担当者が、両手法について差異がないというふうに認識していたところが要因と考えています。

今回、補完手法について確認を行ったところ、サービス統計・企業統計部会と評価分科会での説明に違いが生じているということが判明したことから、改めまして調査研究を行いまして、その結果を御説明させていただくところでございます。

また、今後におきましては、調査手法や推計手法等を変更する際には、管理職を含めて関係職員が情報共有を密に行いながら、内容の理解を深めるなど、十分に注意してまいりたいと考えています。

それでは、資料の方に移らせていただければと思います。

1ページ目を御覧ください。自動車輸送統計調査の概要です。当該調査の目的等ですが、自動車輸送統計、これは自動車による貨物や旅客の輸送実態を明らかにすることを目的とする基幹統計ですが、これを作成しまして、我が国の経済政策、交通政策を策定するための基礎資料を得るために、昭和35年から毎月調査ということで実施しています。

調査の概要ですが、真ん中に表がございますが、様式といたしましては1号様式から4号様式がございまして、1号様式が貨物営業用自動車、これは営業用トラックです。2号

様式が自家用トラック、4号様式がタクシーでして、3号様式が赤枠で線が引いてあるところ、これが今回、御説明させていただきます全数バス調査、これは事業所調査ですが、欠測値補完をしている部分です。3号様式の2と3号様式の3は、それぞれ一般乗合、高速乗合、貸切の自動車の調査ということになっています。

調査方法等と結果公表につきましては、記載のとおりです。

3ページ目を御覧いただければと思います。説明の内容の相違点ということですが、サービス統計・企業統計部会で説明をさせていただいた内容が、左側のロジックAというふうに名前を付けさせていただいております。右側が評価分科会で説明した内容ということで、ロジックBとさせていただいております。現行、調査をやっておりますが、どちらでやっているかということにつきましては、サービス統計・企業統計部会で説明をさせていただいたロジックAのやり方で現在調査を行っているところです。

具体的な内容ですが、まず旅客輸送の特徴として影響を与えるであろうということで、用途と地域と事業規模、これで層分けをしているということです。

まず、用途は、乗合、貸切、特定、車種の区分になります。地域につきましては、地方運輸局別で層を区切っております。

また、事業規模ですが、これにつきましては各用途別の各事業所の実在車両数、これで層を区切っているところです。これにつきましてはロジックA・ロジックB、同じ考え方のもと、層区分を行っているということです。

補完手法ですが、補完手法は4ページ目を御覧いただければと思います。見やすく図式をさせていただきましたが、ロジックAの方をまず、御覧いただきたいのですが、4月分についてはA社、B社、C社、D社の報告がありました。5月分についてはA社、B社、C社の報告があつて、D社については未回収ということで、この条件はロジックA、ロジックB、同じ内容で記載しています。

ロジックAのやり方といたしましては、同一層、先ほど層区分と言いましたが、その同一層の中で事業規模なり用途なり、地域が似通ったところにあるということにつきまして、提出のあつたA、B、Cの4月分と、同じく5月分で提出のあつたA、B、Cそれぞれの層、それぞれの月の走行距離を足していただいて、その足した数字、事業所の層の合計値の前月比を求めます。ここの表でいうと1.3になります。この前月比を4月分のD社の実績値に掛けることによって、5月のD社の欠測値補完の数値を求めるというやり方です。

一方、ロジックBの方ですが、こちらはA、B、Cそれぞれの事業所ごとの4月と5月それぞれの増減率、前月比を出しまして、その前月比の平均を出したものをD社の4月分の実績に掛けることによって、5月分の欠測値補完を行うということです。ですので、層の合計をしたものの前月比なのか、それぞれの事業所ごとの前月比を出したものの平均なのか、ここの違いということです。

3ページ目にお戻りいただきまして、メリットということですが、これは欠測値補完、新しい手法の前のやり方から見たときのメリットということでございまして、これはロジックA・ロジックB、同じなのですが、以前は当該月が欠測した場合、12か月と13か月

前のデータの伸び率を前月の数字に掛けることによって当該月を出していたということでございまして、これをそのまま使っていると、少なくとも1年以上のデータの蓄積が必要であったというところが、このやり方をするによって直近のデータを使うということが出来ますので、1年以内に開業した等の過去のデータの蓄積が十分でない場合にも適用が可能ということと、あとは国内景気等のトレンドとか、イベント等による急激な需要変動を反映できるのがメリットであるというところです。

続きまして、連続補完にする場合のデータの使用期限ということですが、ロジックBで御説明させていただいたときには、連続補完を行った場合のデータの使用期限は最大12か月を目途ということでしてございました。乗合、特定については12か月、貸切については6か月ということでございました。ロジックAの方は特段定めてございません。

続きまして、5ページを御覧いただければと思うのですが、欠測値補完の手法ということで、ロジックAとBの比較検証した内容です。各手法の比較検証につきましては、平成31年3月から令和2年3月の13か月間、この全ての月に提出のあった事業所の調査結果、これは実績値ですが、これを用いまして各手法による推計誤差率及びRMS Eにより行いました。

点線囲みの中に具体的な計算式が書いていますが、推計誤差率ロジックAにつきましては実績値分の推計値ということで、具体的には記載のとおり、未提出と仮定した事業所の実績値分の未提出と仮定した事業所の前月の値×同一層の他の事業所の合計値の前月比、ロジックBの方は合計値の前月比のところ違ってございまして、各事業所の前月比の平均というところに違いがございまして、ここは先ほど説明したとおりです。

RMS Eにつきましては、二乗平均平方根誤差ということで、実際の値からどれだけばらついていくかという数値を見る内容で、計算式は記載のとおりでございまして、推計値－実績値、これに二乗してその層(N)の数分を足して平均を取る。それに平方根を取るというやり方で、誤差を算出するという内容です。

推計値の算出につきましては、平成31年3月分の調査結果を基に、同年4月分の全ての調査結果について1事業所ずつ欠測したと仮定しまして、各ロジックごとに推計をいたしました。また、5月分の推計につきましても、4月分の調査結果に基づきまして推計値を算出し、これを各月12か月分行ったということでございまして、その結果が5ページ、6ページに記載してございまして、用途ごと、ロジックA・Bごとに四分位なり推計誤差を出した表になってございまして、

6ページ目の検証結果でございまして、まず、分かったこととございまして、第1四分位から第3四分位の幅、四分位範囲でございまして、これにつきましてはロジックA・ロジックBで同程度であったということでございまして、また、RMS Eにつきましては、A・Bを比較しまして、ロジックAの方が小さかったということが分かりました。

上記の検証結果によりまして、ロジックAの方が推計誤差が小さいということから、ロジックAが手法としては適切であると考えております。

本調査につきましては全数調査ということでございまして、行政記録情報に基づく毎年の母集団名簿の整備によりまして、事業所の改廃等を適切に把握しているということから、

存在が確認できている事業所の欠測値については、既存データを活用して補完を行うことが望ましいと考えています。そのため、引き続き、ロジック A により欠測値補完を実施することといたしたいということです。

続きまして、7 ページ目を御覧いただければと思います。こちらにつきましては、連続補完によるデータの使用期限の検証というところですが、連続補完によるデータの使用期限の検証につきましては、欠測値補完手法として適切と我々が判断いたしましたロジック A により行いました。

検証につきましては、平成 30 年 3 月から令和 2 年 3 月の 25 か月間の全ての月に提出のあった事業所の調査結果、実績値を用いまして、平成 30 年度と令和元年度、各年度それぞれにおきまして、12 か月連続で補完を行った場合の推計誤差率の平均値を算出し、行いました。

なお、欠測率につきましては、25 か月間の調査期間における各用途の欠測率の最大値を使用しました。

推計誤差率の算出については、以下のとおりですが、平成 30 年度の場合の例を書いています。これは平成 30 年度と令和元年度、期間が違うだけで、やり方は一緒です。

まず、①ですが、平成 30 年 3 月の調査結果を基に、同年 4 月分の全ての調査結果について 1 事業所ずつ欠測したというふうに仮定しまして、全ての事業所の推計値を算出します。

その後、当該推計値をそれぞれの事業所の調査結果（実績値）で割ることで推計誤差率を算出し、4 月分の推計誤差率の平均値を算出します。

5 月分の推計の算出におきましては、今度は 4 月分で算出した推計値を用いて、5 月分の全ての調査結果について、1 事業所ずつ欠測したと仮定し、連続で補完した場合の推計値、推計誤差率、平均値をそれぞれ算出します。当該方法により、12 か月連続補完した場合の推計誤差率の平均値を算出したということです。

各月の推計誤差率の伸び率に、上記で御説明しました 25 か月の調査期間における各用途の欠測率の最大値を乗じ、各用途の欠測率における推計誤差率の平均値を算出したということです。

その算出結果が 8 ページです。平成 30 年度、令和元年度と、それぞれ 12 か月連続で欠測したものであるということになります。乗合は、欠測率 17.1% も欠測率で算出した内容となっています。貸切につきましては欠測率 20.1%、特定については欠測率 13.9% ということで計算をした結果です。

検証結果です。9 ページを御覧いただければと思いますが、25 か月の調査期間における最大の欠測率を用いまして 1 年間連続補完したとしても、各年度とも大幅な精度の低下は見られないと考えています。

令和 2 年度における平均欠測率につきましては、乗合が 9.0%、貸切が 11.4%、特定が 7.6% ということですので、検証で使用した欠測率よりも更に小さいことが、割合的にも実際の推計誤差率もより小さくなるのではないかと想定されます。

また、検証に用いた欠測率は、同じ事業者が連続して欠測することが前提となっておりますが、実際には欠測となる標本が入れ替わるということになりますので、誤差はより

小さくなるということが想定されます。

また、令和2年度の調査票の平均回収率が80%以上ということでございますので、欠測値補完による推計値が統計全体の集計に与える影響というのも小さいのではないかと考えてございます。

上記を踏まえまして、輸送実績を適切に捉えるために、連続補完によるデータの使用期限というものは定めず、前年同期を超えた場合も引き続きロジックAの手法による欠測値補完を行いたいというふうに結論付けております。

ただし、連続補完を行うに当たっては、より一層の督促の強化を行うということと、より適切な母集団の把握、具体的には毎年母集団名簿を整備していますので、この整備に加えて、毎月の調査時で把握しました廃業とか実質廃業とか、そういう情報を速やかに反映するというところを行いまして、補完が不適切な対象を減らしていくということと、連続補完の期間が短くなるように努めてまいりたいと考えてございます。

10ページからは参考資料ということですが、10ページはサービス統計・企業統計部会で御説明させていただいた資料を参考に付けております。11ページから最終ページの13ページまでが、評価分科会で説明させていただいた資料です。

国土交通省からの説明は以上です。御審議よろしくお願いいたします。

**○菅分科会長** ありがとうございます。それでは、ただ今の御説明について御質問、御意見を賜りたいと思います。手元に出席者一覧の名簿がありますので、その順番で委員の先生方の御意見、御質問をお聞きしたいと思います。

まず最初に、川崎委員、御質問、御意見を賜りたいと思います。

**○川崎分科会長代理** この件、過去には審議されたようですが、私としてはその当時のことをよく知りませんし、また今日の会議の趣旨が、ただコメントを言うだけでいいのかどうかよく分からないので、状況が分からないなりに質問や感想を申し上げます。

欠測値の補完ということですが、まず体系を正しく理解しておきたいので、質問の1点目は、先ほどのお話ですと、欠測の発生がものすごく大きいわけではないという印象を受けました。資料の何ページかちょっと見当たりませんが、10%前後かそういう数字が出ていたように思いますので、そんなに大きい数字ではなかったように思うので、補完の精度を上げれば上げたで、それはそれでいいと思いますが、物すごくこれがクリティカルな要素になるとはあまり思わないので、これぐらい緻密に検討されていけばいいのではないかと考えたのですが、ただ1点だけ気になったのが、連続補完という概念がよく分からなかったということです。

そこでちょっと教えてほしいのですが、2か月以上続けて欠測になるのがどれぐらいあるだろうか。それは、例えば1年ぐらい続くのがあるのだろうか、そういう欠測の連続の分布みたいなデータというものはあるのだろうかというのが1点お尋ねしたいことです。これは連続補完ですね。

それで、その場合に連続補完というか、例えばこの方式だと、前月の比率をずっと掛けながらやっていくということになるのだと思うのですが、例えば3か月欠測が続いているところであれば、1か月目の補正率と2か月目の補正率と3か月目の補正率全部を掛けて、

最後の観測値に掛け算するというやり方なのか、どういうやり方をされているのかという実態がよく分からなかったので、そこを教えていただきたいというのが1点目です。欠測の発生と連続補完の考え方を教えていただきたいというのが1点目です。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 御質問ありがとうございます。分布図というのを御用意してないのですが、連続補完がどれくらいかということについては、各欠測の半分ぐらいが1年以上提出されていないという内容になります。

○川崎分科会長代理 ちょっと確認させていただきますが、資料の9ページ目の1番目の資料、2番目の資料もあるんですね、それに令和2年度の平均欠測率が10%前後、乗合9、貸切11、特定7と出ているので、これは単月当たりの欠測値ですね。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 はい。平均欠測率は、令和2年度における欠測率を平均したものになります。

○川崎分科会長代理 そうすると、ある月、特定の月を見ると、例えば100事業所があったら、9事業所が欠測になっているということなのだろうと思うのです。私がお尋ねしたい趣旨は、9の事業所が、例えば翌月2か月続けて欠測になるのがどれくらいあるか、翌々月の3か月続けてなるのがどれくらいあるのかという趣旨なのです。それが全部ばらばら入れ替わりながら、ランダムに入れ替わるような感じなのか。それとも、特定の9%がずっと続いているのかという趣旨なのですが、その辺りはおよその感覚的なことでもいいのですが、分かりますでしょうか。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 平均欠測率は大体9%、11%と申しましたが、その半分ぐらいが固定してというか、毎回欠測しています。

○川崎分科会長代理 だから、例えば9%なら、四、五%ぐらいが数か月つながって欠測になるということですか。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 さようです。

○川崎分科会長代理 分かりました。そういうことですね。ただ、そうすると100%に対しての四、五%とか、それぐらいの感じだから、あんまり影響は大きくないなというのを直感的に感じたところで、その点は数字がそんなに大きくなさそうというのを聞いて、多少安心したところでは。

そこで次にお尋ねしたいのですが、これは推計方法の問題よりもデータの利用について教えていただきたいと思います。この結果数字というのがどれくらい厳密な利用を求められているかによって、欠測値の補完をどれくらい厳密にやるかというのは変わってきていると思っています。

例えば個別の給付額に連動するとか、例えば過去の例でいえば、毎月勤労統計みたいに雇用保険の保険料に連動するようになってくると、相当厳密にやっていかないといけないところがあるのですが、例えばマクロ経済の動向を把握するとか、およそ経済が上向いているか、下向いているか、それはスピードがどれくらいあるか、そういう感覚を捉えるぐらいだったら、ある程度の誤差は容認できると思うので、そういった利用が、例えば給付や再配分にリンクするような指標があるのか、あるいはそれはないけども、マクロ経済の分析に利用されるか、どういうところで利用されるか、あるいは数字の規模を捉えるために利

用するのか、変化を利用するのか、そのような利用の状況は分かりますでしょうか。そこから辺を考えながら、この議論をどれだけ厳密にするかと考えた方が私はいいように思っているのですが、その利用の実態など、幾つか例などを交えて教えていただいたらありがたいのですが。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 ありがとうございます。当該数字については、予算とか予算に係る指標とか、そういうものに使われているということはお聞きしておりません。どちらかというところ、それぞれの業種について当該輸送状況、輸送量がどういう状況にあるのかということとか、あとは実際モーダルシフトになったときに貨物はどういう状況になるのかとか、そういうところの指標ということで使われるのですが、具体的に何かの交付金とか、予算のために数字を厳格に見るとどうこうという指標にはされていないということです。

○川崎分科会長代理 分かりました。ということは、現在されているようなこのような補完の方法というのは、精緻化すれば幾らでも精緻化の余地はあると思うのですが、それぐらいの程度のことを考えておられるんだとしたら、この目的にはかなった方法ではないかなというのが私の感想です。

その上で、もう2つだけ、これも感想として申し上げれば、できるかどうかは別として、私が直感的にやったらいいのかなと思ったのは、平均よりも変化率の中央値、個別の事業者ごとの変化率の中央値とか最頻値を使うのが本当はいいのではないかなと思います。ここではロジックA・ロジックBとおっしゃっていますが、私の話はロジックCとかDとか言っているような話ですが、今お見せになっているスライドのもう一つ前にポンチ絵のようなものがありましたけれども、これでいうと、ロジックBで単純平均を取ると、これは極端に高い値、低い値が出てくる、それも平均を取るわけですが、私はむしろ欠測が起きているところに対しては、一番ありそうなのは中央値だと考えるのがいいだろうと思うので、たくさんある中の中央値を取るとか、あるいは分布がある程度きれいな形をしていれば、最頻値を取るのが本当はいいのではないかなと思います。ただ、恐らくそれをやっても結果は大きく変わらないだろうから、これはあくまでも感想で、やってみるまでのことでもないのかなとも思っています。これが一つ感想です。

それから、あともう一つ、これも感想で、できたらいいと思うことを申し上げますと、補完に使う比率をずっと固定していくというお考えですが、計算すれば毎月出せるのに、どうして毎月出されないのか。毎月出して、それを使っていかれるのが一番フレッシュでいいだろう、毎月毎月更新していくというのが本当はベストだろうと思うので、固定するのはあまり賛成できないと思います。ただ、何らかの事情があるのか、それは計算システムとか、そういうところの制約があるのかとも思いますので、もしそうならばもっと頻度を上げられるように工夫していただくのが重要じゃないかなと思います。これも感想です。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 比率を固定というところは、連続補完するに当たっては、毎月毎月補完のための推計値が出てきますので、その推計値を基に、あとは実際に出てきた当該月のその層の事業所の比率なりを使って推計した数字を使っていきますので、そういう意味では比率は固定というよりは、毎月毎月という

こととございます。

○川崎分科会長代理 わかりました。9ページ目の資料の検証結果で、連続補完によるデータの使用期限は定めずと書いてあるので、何か同じものをずっと固定していくという意味かなと思いましたが、これは毎月毎月、計算し直しながら、延長していくという意味なのでですね。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 さようです。それをすると、その結果が、8ページにあるとおり、推計誤差率が大体1.0何とかという、比較的低い数字で推移するということです。

○川崎分科会長代理 ここの使用期限は定めずという言葉で何か固定されるというふうに私は誤解をしてしまったのかもしれないので、その点もう少し分かりやすく教えていただけたらと思いますが。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 評価分科会で以前ロジックBを御説明したときは、使用期限を最大12か月目途というふうに期限を定めておりました。それに対して、今回は使用期限を定めずというところです。

○川崎分科会長代理 分かりました。

○菅分科会長 神林委員の方から御意見、御質問を賜りたいと思います。よろしくお願いたします。

○神林臨時委員 ロジックAとロジックBの違いについてなんですが、平均を取って変化率を取ると、変化率の平均を取るのではばらつきが異なるというのはメカニカルに違いますので、どちらがよいのかというのを評価するのは、それほどアドホックにできる話ではないと思います。

この統計を取るというか、報告する目的が各月の総距離数を計算することにあるのか、それとも伸び率を計算することにあるのか、どちらを優先して統計を作るのかと関係すると思いますので、利用する数字の目的に応じて補完の仕方を選んでいくのが正常なやり方だと思います。その利用目的を特定せずに、ロジックAとロジックBを比較して、ばらつきがどちらが大きいという話で選択する話ではないかなと思います。それが1点。

もう1点は連続補完の話なのですが、これは前提として監査データがあるのですね。それで閉鎖した事業所に関しては、廃業がわかりますので、これは母集団のところ、9ページの最後のところですが、廃業した事業所については母集団名簿の方からアップデートできる。ですので、欠損で補完する事業所は、母集団名簿からは明らかに営業している事業所ですが、調査票を出してくれていない、回答をしてきていないのが確定している事業所ということになるわけですね。

そういう実際に営業しているのだけれども、調査票を出してくれない事業所があったときに、その状態が12か月続こうと25か月続こうと、しようがない、やっていることは間違いないけど、調査票は出してくれない、だから補完をし続けましょうという判断だということになるのではないかと思います。ほかの統計類ですと、母集団の整備がこれほど頻繁に行われませんので、廃業してしまったから調査票が戻ってこないのか、それともただ回答拒否を続けているから、調査票が戻ってこないのかが分かりません。ですので、ある

程度 12 か月とか6 か月とか、連続して調査票が戻ってこなければ、とりあえず廃業したのではないかと推測して、サンプルから外すということをやっているのだと思います。

その違いがこの統計の場合にはありますので、調査母集団は行政データで押さえられている、やっているかどうかというのは分かっているということが多分最初の方にちゃんと書いた方がよいのではないかなと思います。

それでしたら、欠損を補完し続けるのは仕方がないのかもしれないのですが、何ゆえこんな 25 か月とか出してくれないのかということは、きちんと別途検証すべきだろうと思います。つまり、ランダムに出してくれないのであれば平均値でカバーする、あるいは中央値でカバーするということは多分理論的にも正しいのですけれど、25 か月連続でやっているにもかかわらず絶対出してくれないのは何かある。背後にその要素というか、要因があるということになると、平均値でリカバーして構わないのか、少し疑問が出てくると思います。

何らかの特殊な事情があるので、25 か月連続で出してくれないということを考えると、平均値でリカバーするよりは、例えばもっと裾野あたりというのでリカバーした方がいいのかもかもしれませんし、それは別途、特にケーススタディーのような形で少しずつきちんと情報を集めていって、25 か月連続、12 か月連続欠損を超えた場合には調査員を派遣して、どうしてこういう状態なのかというのをきちんと把握するというのを併せてやった方がよいと思います。多分やっているのではないかと思いますけれども、以上2点です。

#### ○内田国土交通省総合政策局情報政策課交通経済統計調査室長

まず、どのような指標で使われているかということですが、自動車輸送統計は基本的に輸送の大きさをメインで見えておりますので、今回のバスの関係ですと、輸送人員とあと輸送人キロという視点で見ることがほとんどです。

また、25 か月以上とか、12 か月以上出していただけないところですが、調査対象が全事業所で4,400 ぐらいです。そのうちの8割以上は回収できていて、残りの部分が未回収でして、基本的に毎月督促をしています。母集団名簿はあるのですが、廃業届とかを出していただくタイミングが遅かったりしますので、その中でも調査票が戻ってきてしまい電話をかけて連絡が取れないとか、あとは実際にホームページとかも調べて載っていないとか、そのような状況や実質廃業という情報を見極め、それを母集団名簿に反映しています。

○菅分科会長 久我委員いかがでしょうか。御質問、御意見ありましたら。

○久我臨時委員 御説明でロジックA・Bの比較検証ですとか、連続補完によるデータ期限の検証などをしていただいて、いずれも検証の結果、実際に生じている誤差よりも小さく判断できるということで、私の方から方向性について特に異論はありません。また、川崎先生、神林先生の御質問で理解も進んだところです。

細かいところで、念のため単純な質問があるのですが、資料の9ページ、2点目と4点目のところですが、先ほど川崎先生が乗合9.0%とか、この辺りを御質問されたのですが、これも令和2年度における平均欠測率ということですが、過去も大体同程度なのか、あるいは上昇傾向にあるのかないのか、その辺り、この程度なので、実際の検証結果よりもこちらが小さいのかという御判断だったので、数値のトレンドというのを教えてい

ただきたいというのが1点と、同じように御判断の材料として4点目、調査票の回収率が8割以上と高いのでという御判断なのですが、こちらもずっと同程度を維持しているのか、そのトレンドがどうなっているのか、この2点教えていただけますでしょうか。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 調査票の回収率は、検証した25か月のときに比べて督促等をかなり頑張って実施していますので、以前は80%を下回っていましたが、令和2年度の平均は80%以上で、回収率は上昇傾向です。それに伴いまして、今まで出していただけなかったところからもいただけるようになっていまして、欠測率は減少方向です。

実際事業をやっている今出てないところに、より一層督促を強化し、欠測率を下げていくということと、そもそも1回も出してくれてないというか新たに調査対象となったが出してくれないところは、一回は出していただき補完により数字に取り込んでいきたいということです。引き続き回収率の強化というのはしていきたいと思っておりますし、それに伴い欠測率は今後更に下がっていくのではないかと考えています。

○久我臨時委員 今回の判断の根拠としている数値は改善傾向にあって、更に誤差は実際には減るのではないかとということで承知いたしました。

○菅分科会長 山本委員の方から御質問、御意見よろしくお願ひいたします。

○山本臨時委員 およそロジックBではなくて、ロジックAを使うということは、私も先生方と同様、悪いことではないだろうという感触を抱いております。

その上で、質問です。4ページに補完方法がございまして、ロジックAで比率を計算する際の様々な調査値の合計は、4月と5月の両方に存在、回答している事業所の場合は、4月と5月の両方、回答している事業所のみ合計ということによろしいのでしょうか、それとも回答したものの全部合計しているのでしょうか。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 層区分した中で、4月・5月両方、回答があったところの合計です。

○山本臨時委員 おそらく、基本は伸び率補完というやり方ですね。

それから、コメントが幾つかあります。先生方が気にされていますように、検証結果の10ページぐらいのスライドで今御説明いただいていた、乗合が9%、貸切が11%、特定が7.6%の欠測が大体ある中で、回収率は全体80を超えておりますが、それらの半数が1年以上欠測していると御説明いただいたのですが、連続欠測回数の分布があると少し説得力が増すような気がしました。

最初、川崎先生の御質問に対して、半数程度が1年を超えるという御説明をいただいたような気がしますが、特に連続補完によるデータの使用期限は定めないということに関して、最大でどの程度長く使うことになるのかということがやはり気になることは気になります。

ですので、およそどれぐらいの長さの分布になっているのか、月数とか回数でいいと思うのですが、連続でどれぐらい長くなるのかという分布が気になります。また、回収率で見ますと、統計の規模がちよっと分からないので、例えば走行キロ数や輸送人員に関して、大きな値を持っている事業所を少ないが落としているのか、それとも小さな規模の事業所

が多くなっているの、合計を推計している本統計としては先ほど御紹介いただきました試算結果のように問題がないということなのか。その辺りを今回使用期限を定めないことに関して御検討いただいてなくて、ロジックの上のところ、4ポツは全部、BよりAがいいということだけの報告だと思えますので、総論としての方向に私も異議はございませんが、念のため本当に定めなくて大丈夫なのかということについて、内部でも構いませんので、引き続き御検討いただけたらと思いました。これはコメントです。

○菅分科会長 恐らく後半はコメントですので、それを今後御検討いただくということで、前半については両方だということで、御回答いただけますでしょうか。前半については、両方とも回答した事業所について比率を取っているという回答でよろしいでしょうか。

○山本臨時委員 私もそう御回答いただいたと承知しております。

あとは、追加で申し訳ないですが、6ページの伸び率補完してみたという、その辺りの結果について、乗合はいいのですが、中央値がほぼ1というのがよく分かりませんでした。それはちょっと置いておくとしても、平均値がおよそ0.5増しになっていますので、平均とかRMS Eについて計算すると、そもそもこの補完の仕方は結構大きい方に偏っていることになりやすいですね。

気になりましたのは、平均値の方が第3四分位よりも大きいことが多く、とても大きい値が結構出て引っ張っているのではないかと思います。第3四分位まで超えることはたとえば正規分布だとあまりないように感じます。それでRMS Eを比較されているのですが、これはもしかしたら偏りと分散に分けて検討するのが本来よかったのかもしれないという印象を抱きました。これはコメントです。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 御指摘いただいて、引き続き見ていきたいなと感じたところです。偏りと分散、コメントとして承りまして、何か分かることがあれば、我々の中でももう少し調べられれば調べていきたいと思えます。

○菅分科会長 これについては、再度、国土交通省さんの方でちょっと検討していただいた方がいいと思えます。これで大丈夫かなということ。

○山本臨時委員 おそらく偏りがとてもある分布なので、平均で平均を当てるのは難しい状況だと思います。例えばどちらが望ましいという議論をする意図は私にはなくて、どちらもそれほどよくないかもしれません。その中で悪くない方を取っているとすると、結論は一緒になります。そういう意味で少し御検討いただきたいという程度です。この資料そのものを見て、どっちもよくないのではないかと、他を考えなさいという意図は私の発言にはありません。

○菅分科会長 では、そのようによろしく願いいたします。

以上、先生方の御質問、御意見を賜りました。川崎委員の方からは、それほど大きくないのではないかと、クリティカルではないのではないかと。連続補完というのがちょっと分かりづらいということと、具体的に連続補完の分布がどうなっているのかということが重要ではないかということで、これについては山本委員からもそれがあつた方が説得力があるのではないかとということでした。

もう一つ、川崎委員から御指摘がありましたのは、データの利用との関係が重要で、ど

れくらいの厳密さが求められるのかということで、これに関しては交付金給付とのリンクがないので、それほど厳密性は今のところ求められていないということでした。

また、神林委員からは、AとBのどちらがいいのかを考えると、アドホックではなくて目的に応じて考えるべきで、レベルなのか、それとも変化率なのかということで、これについてはレベルであるという回答でした。また、廃業について行政記録で把握しているということだが、連続的に回答がないのであれば、それについてはよく調べるべきではないかという御意見でした。

久我委員からは、欠測率に関して令和2年だけ載っているが、過去はどのようなかということ質問があり、国交省よりは過去は80%よりも更に低かったという回答があり、これについては改善傾向にあるということは評価できるということでした。

山本委員から質問としては、そもそもこの比率は両方回答した、要するに継続標本で計算しているのか、そうでないのかという質問があり、国交省からは継続標本であるという回答があり、また、分布の特性についてもう少し検討、特に大きな事業所の値等を検討すべきではないかというご意見がありました。

**○川崎分科会長代理** 1点だけ、神林さんがおっしゃったことはすごく大事なので、改めて私の方からも強調してお願いをしたいことがあるので、申し上げます。

それは廃業の確認です。というのは、これは既にお話しになったことですが、廃業されたところを誤って補完をやってしまうと、総数が伸びてしまいます。一体いつ廃業したのかというのが実は結構問題で、廃業した、していないというところで1事業者分、過去の数字がそのまま載る載らないでゼロになるのは、大きな違いが出てくるわけですので、ここは相当丁寧にされないともうまいだろーと思えます。

これは許可業種ですので、登録を抹消したら廃業かという、私はひょっとしたらそうでないのではという気がしています。調査項目は企業の業績のようなものですから、徐々に売上げが、あるいは顧客が減少しているところは長期低落傾向があって、あるところで極めてゼロに近づく、あるいはゼロだけでも、登録抹消もしないということが起こっているかもしれない。しかもそこに未回答が重なると、これは過去の数字をそのまま載せたことになるので、補完すると、過大になってしまいます。

そういう意味で、未回答の事業所に対してそのような長期低落傾向がないかとか、少しパネルデータ的にフォローして、そのような閉業間際のもを誤って補完するということのないように、チェックはしっかりしていただかないと、これは総数に響くのではないかと思いますので、その辺を是非頑張ってくださいというのがお願いです。

**○菅分科会長** これは廃業確認等について、休業との違いが微妙なのですが、それについてちょっと慎重に、単に補完するのではなくてということだと思います。

**○内田国土交通省総合政策局情報政策課交通経済統計調査室長** 毎月督促という行為は、4,400事業者の一部でございます、督促はしておりますので、そのときに状況、実際やってなければ、やってないから送らないでくださいとか、そういう話はいただきますので、その中でおおよそ実態は把握できるのかなと思っております。

**○神林臨時委員** この統計のとりあえずの目的が総キロ数であったとすると、例えばロジ

ックAを使うときに、今、山本委員の質問に対する回答として4月と5月、あるいは両方存在した事業所に関してのみ使って、補完するための数字を作ることになっていると思うのですが、4ページの図を見ると分かりやすいですね。総距離数であれば、4月に存在して、5月に廃業してしまったという事業所も、4月には走っていたわけですよ。逆に5月に新しく開業した事業所も5月に走っているわけですので、その情報は使わずに、補完数字を作ることになります。

こういうバランスしたパネルで作るのであれば、ロジックAとロジックBはそれぞれのレベルのシェアと伸び率がどう相関するかによって答えが変わってくるだけなのですが、ロジックAを使うメリットのもう一つは開業した事業所、廃業した事業所の情報を使えるということだと思います。こういう事業所は伸び率は計算できませんから。ですので、ロジックBでは使えないのですが、ロジックAでは使える。

回答してくれた事業所の総距離数を各月で計算をして、その比率を取って、回答がなかった事業所の伸び率というのをそこから推計してあげるといふ計算の仕方もあり得ると思います。そちらの方が情報としては全部使っていますので、統計的な精度というよりはロジカルな精度が上がっていると思います。使える情報を使っていますので。ですので、ここはほかの先生方の御意見も交えて検討していただければと思います。

○菅分科会長 要するに片方が存在しなかったら、次の月で存在した場合どうするか、あるいは逆のケースですね。

○神林臨時委員 似たような企業、似たような走向距離数の事業所が出たり入ったりしていれば、結局倍数にしていますので、関係がない。違うかな、全部総足しているから、ちょっと関係しますね。ただ、入ってきた事業所はおそらく大きいですよ。廃業する事業所は、おそらく、廃業直前はあまり走ってないですね。というようなことを考えると、総距離数を計算するときに両月に存在している事業所だけで計算するのと、単月にそれぞれ存在した事業所の総距離数を計算するのでは結果が異なってきますので、後者を計算するとどれだけ変わるのかというのをきちんとフォローしておくべきだろうと思います、ほかの先生方もし御意見があればお聞きしたいのですが。

○菅分科会長 確かにおっしゃるとおり、前の月はなかった、今回入ってきたという情報もその次の月から入ってくるので、最初の1か月は稼働がそれほど、例えば1日から稼働していれば確かにそうだけでも、例えば月の半ばで稼働すると半分ぐらいなので、そこはちょっと処理が難しいかなとは思いますが、どれくらい参入・退出があるのでしょうか。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 具体的な数字というのは持ち合わせてないのですが、乗合とかはそれほど出入りがあるということは考えにくいかなと思っています。こういうコロナ禍でありますので、参入が多いかということとはなくて、廃業の方が多いというのはまず言えます。乗合については基本的に乗合バス、町なかで走っている事業所でございますので、そういう意味ではすぐ廃業という話にはなりにくいということです。貸切については、世の中のトレンドによって増えたりとか、コロナ禍だと実質廃業というか、動かないとか、そういうことがありますので、貸切については、今の状況ですと、廃業の方向での出は多いのかなと考えています。

○菅分科会長 参入の方は、貸切の方は出入りがあるので、その影響はもう少し検討した方がいい感じはありますね。要するに無回答の場合、それはどういう状況なのかという、それは本当に廃業と言えるのか、単なる休業なのかとか、特にコロナなどの場合、あり得ますね。それを単純に補完していいのかということと、その辺りの情報をちょっと検討する余地はあるのかもしれないということですね。

つまり、参入・退出の状況に関するデータがないと何とも言えないところがありますが、検討した方が、要するに乗合に関してはまずあまりないと思いますが、貸切は確かにあり得るので、それはちょっと慎重に、例えばコロナなどのときはどうなっているのか、それを単純に欠測値地として補完していいのかというのは少し疑問の余地はありますね。

神林委員。今の議論なのですが。特に答えがあるわけではないのですが、参入・退出の状況に関する情報がないと議論しづらいところはあると思いますけれど、乗合に関してはあまりない。貸切についてはかなりあり得ると。そこについてもう少し参入・退出状況を見て、ほかの方法を検討した方がいいとは言えると思いますけど。

○神林臨時委員 そのときに、参入・退出の件もそうなのですが、ターゲットにしているのが総距離数で、退出した事業所が退出する前に持っていた距離が問題ですので、事業所の数で何%とかいうのではなくて、500キロ分いなくなりましたとか、800キロ分いなくなりましたとか、500キロ分新たに入ってきましたとかいう、その数字でのシェアがもっとクリティカルになるのではないかなと思います。

いずれにしても回答率が8割超えていますので、総計してしまうとすごい数のストックが既に計算されていますから、よっぽどすごい事業所が出たり入ったりしない限りは、総計はそう簡単に動かないと思いますので、計算してみたら微妙に変わるだけだとは思いますが。けれども、方法として正しい方法を考えるべきだろうと思います。ロジックAであれば、存在した事業所を全部総計する方が自分は正しいのではないかと思います。山本委員、どうでしょうか。

○山本臨時委員 私に特に答えは持ち合わせておりませんが、1事業所当たりを基準化して伸び率を計算するのと、先生おっしゃいましたように市場規模なので、合計で計算するのとどっちがいいかなというぐらいの悩みで、今の話を伺っていました。しかし、もし今後、補完方法についてさらに検討していただく、つまり言い換えると、改善していただく余地があるとすれば、単に両方ある共通事業所の伸び率だけではないやり方もあるような気はしております。

○川崎分科会長代理 今回の御議論は大変参考になりました。私は、どちらかと市場規模に連動するというのが原則としていいと思っています。そういう意味で全部合計同士での比率でいいのではないかと思います。ただしそれにはやっぱり前提があって、要するに無回答という現象が業績と連動しているかどうかというのが問題だと思ひ、あす。結局、業績低落中のところにそれを全体の平均と同じように伸ばしたら、これはやり過ぎになるわけですね。

ですから、先ほど神林さんもおっしゃったように、無回答が全くランダムに発生して、消えてはまた戻ってくるみたいな現象が起こっているときであれば、私は今のような方式

でいいだろうと思うのですが、ただ分析してみて、明らかに低落ぎみの事業者ばかりが無回答だというときには、総計の比率方式はやり過ぎになる可能性はあると思うのです。

ですので、むしろ国土交通省の方々をお願いしたいのは、今すぐということではなく、これはちょっと分析に時間かかることだと思うのですが、例えばこれ4,000ぐらいの事業所なので、データをパネル化するのは多分それほど難しくないだろうと思うので、一度パネル化して、欠測の発生とか戻り方とか、あるいは業績のアップダウンのトレンドみたいなものを少し見ていただくとか、そういう分析を少ししていただいた上で結論を出していくのがいいのではないかと思うので、非常に大事な指摘を今神林委員にさせていただきましたので、そこはデータがない以上はこの議論だけで結論が出ないということだと思いたすので、そういうことをしていただけたらと感じました。

○菅分科会長 いかがでしょうか。継続標本だけではなくて、比率を取るべきだという選択肢もあるということで、それについてはパネル化して、実際に実験してみないと何とも言えないということで、今後御検討いただけたらということでもよろしいでしょうか。

○内田国土交通省総合政策局情報政策課交通経済統計調査室長 持ち帰らせていただいて、検討をまずさせていただければと思います。

○菅分科会長 そうですね。いかがでしょうか。そのほかに議論ありますでしょうか。御意見等。

○川崎分科会長代理 この議論、どこで終止符を打つか、コンクルージョンするかというのは非常に難しいと私は思うんですが、正直言いますと、補完の推計というのは改善していこうと思えばいろいろな改善の余地があるわけで、これで終わりということは多分ないのかもしれないです。

ただ、だからといっていつまでも延々とやっていくのも、これまたリソースを幾らでも食ってしまうことになるので、程々でやらなければいけないと思うのです。ですので、私自身は今のロジックAを採用するということと、それから特に山本委員がおっしゃったような期限なしにやっていくのが本当にいいのかというのはちょっとペンディングだということにしつつ、そのうえで、大きな結論として私はこれで了解ということで、ただ今後、国土交通省御自身の検討課題として、今出たようなことを更に改善の余地はないかということを見直していただくことが大事ではないかと思いたすので、これは欲を言えば、どんどん精度を上げていく方向はまだまだあり得ると思うので、そこはエンドレスに、私自身この場で国土交通省に求めていくというほどのことではなく、むしろ御自身の努力で頑張ってくださいたいと感じたということです。

○菅分科会長 まとめていただきまして、ありがとうございます。私もこれはちょっと限界があるだろうとは、もともと補完ということ自体に限界があるということと、2つの方法でそれほど大きな違いが出るはずがないというのも思いたすので、ロジックAで継続していただいて、ただし今回のケースに関しては、当初の評価分科会の説明と実際の運用が違ったというのは問題だったと思う。むしろそちらの方が問題で、意思疎通ですね、何をやっているのかということに関してこういうことがないように十分をお願いいたします。

川崎委員からまとめていただいた形ですけれども、本件に係る国土交通省の対応に関し

て、大きな方向性としては問題ないという形だろうと思いますが、考え方としてはより整理すべきところがあるような御指摘をいただいたと思いますが、おおむね妥当であると評価したいと思います。そのような形でよろしいでしょうか。特に問題ないでしょうか。

なお、今回の検証の中で示された数値について、念には念を入れて、誤り等がないかももう一度確認するよう、国土交通省さんにはよろしくお願ひしたいと思ひます。

大変有意義な議論だったと思ひます。予定されました議事が終了しましたので、本日の分科会はここまでとさせていただきます。

最後に、事務局から御連絡事項をお願ひします。

○**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、日程調整をさせていただきますと思ひております。

○**菅分科会長** 本日はこれまでといたします。皆様には大変お忙しいところ、審議に御協力いただきましてありがとうございます。感謝いたします。ありがとうございました。

それでは、これにて終了したいと思ひます。