

資料 1

保存期間：5年
(令和8事務年度末)
令和3年11月19日

民間給与実態統計調査の 復元推計手法について

1 はじめに

民間給与実態統計調査においては、その復元推計手法について、第1回検討会にて、行政記録情報の活用等を含め、類型ごとに課題の整理を行った上で、対応方法を検討した。

今般、既に検討した内容に加え、「母集団サイズ」及び「調査対象外事業所の推定方法」について、次ページ以降のとおり、更なる精度向上を目的として、キャリブレーションの手法等を活用した追加的な検討・整理を行った。

○第1回検討会にて検討した復元推計手法の見直しに係る項目等

「母集団名簿の整備」

・よりの確な母集団サイズを推計するため、標本事業所に含まれる調査対象外事業所について、復元推計へ適切に反映させる。

「階層が異なることとなった事業所への対応」

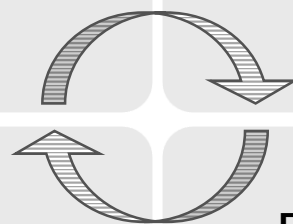
・抽出率の逆数を使用した復元推計の実施に当たっては、全ての事業所について、標本事業所として抽出した際の抽出率を使用した計算を行う。

「従来の統計作成手法の見直し」

・労働力調査をベンチマークとした従来の処理を廃止する。

「無回答の処理方法による過大復元の是正」

・低階層の事業所を中心とした未回収の発生割合に係る有意な差を補正するため、税務データを副次的情報として活用等した計算を行う。



2 母集団サイズについて


(1) 課題

第1回検討会においては、「母集団名簿の整備」として、よりの確な母集団サイズを推計するため、標本事業所に含まれる調査対象外事業所について、復元推計へ適切に反映させる手法を検討したところ。

一方で、民間給与実態統計調査における母集団サイズについては、当該調査対象外事業所を要因とするほか、標本抽出時から調査の基準日（12月末）に向けて、給与支給人員が変動する事業所等が存在することによって、枠母集団サイズと調査の基準日におけるターゲットとすべき母集団サイズとの間に、一定の開差が生じる形となっている。

そのため、よりの確な母集団サイズを推計するに当たっては、ターゲットとすべき母集団サイズへ向けた復元推計となるよう、調査対象外事業所を加味する手法に加え、追加的な対応方法を検討する必要がある。

【枠母集団サイズと調査の基準日における母集団サイズが異なることとなるイメージ】

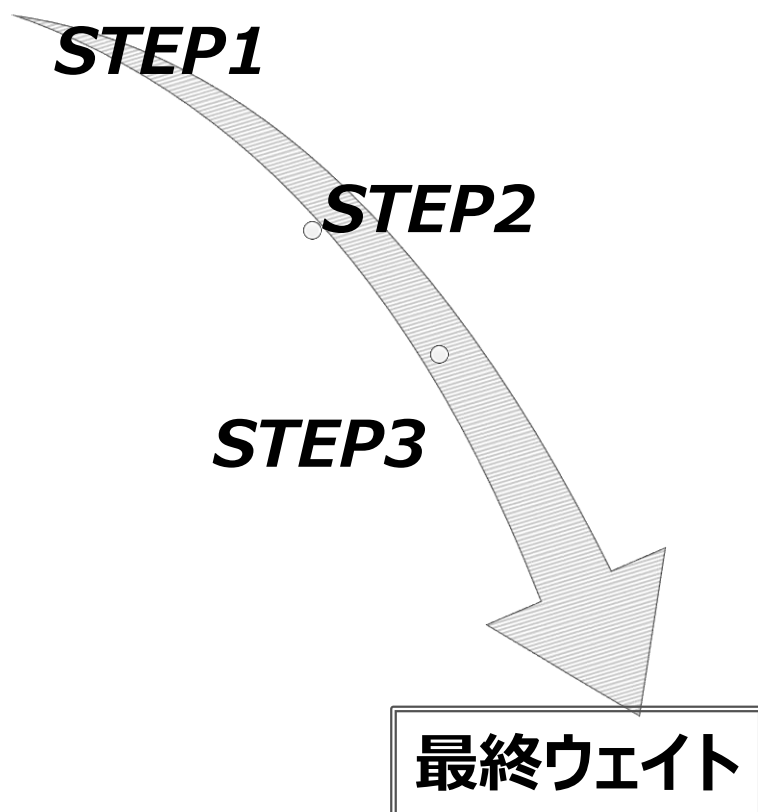
		調査対象年								
		...	8月	9月	10月	11月	12月	1月	2月	...
処理内容等	● 標本抽出に使用する枠母集団サイズの作成							● 調査の基準日		
	<div style="border: 1px solid black; padding: 10px; width: fit-content;"> 枠母集団サイズ 1層_100件 2層_80件 3層_70件 4層_60件 5層_50件 ⋮ </div>							<div style="border: 1px solid black; padding: 10px; width: fit-content;"> 母集団サイズ 1層_80件(100-20) 2層_85件(80+5) 3層_65件(70-5) 4層_60件 5層_50件 ⋮ </div>		
		≪例示≫ ・調査対象外事業所が20件 (全て1層の事業所とする。) ・ <u>3層から2層へ変わった事業所が5件</u> 存在した場合								

2 母集団サイズについて

(2) 対応方法の検討

母集団サイズへ復元推計する過程において、抽出率や回収率の逆数を乗じるほか、ターゲットとすべき母集団サイズへ調整する処理（キャリブレーション）を実装することで対応することとしたい。

具体的には、以下のとおり、計算することとしたい。



$$\text{最終ウェイト}^{\ast} = \text{STEP 1で求めた乗率} \times \text{STEP 2で求めた乗率} \times \text{STEP 3で求めた乗率}$$

※最終ウェイトとは、母集団サイズへ復元推計する際に使用する乗率

- **STEP1（標本として選ばれる確率の逆数について）**
標本として選ばれる確率（抽出ウェイト・基礎ウェイト）としては、「抽出率」を使用する。
- **STEP 2（回収になる確率の逆数について）**
回収になる確率としては、「抽出層（階級単位）ごとに算出した回収率」を使用する。
- **STEP 3（キャリブレーションの実施について）**
別途、調査の基準日におけるターゲットとすべき母集団サイズを算出した上で、当該母集団サイズに、STEP 1とSTEP 2で求めた乗率を掛け合わせたものの総和を合わせる作業を行う。
さらに、税務データを使用して欠測値補完を実施した上で、復元推計した給与支給人員をベンチマークとした計算を行う。

2 母集団サイズについて

(3) 具体的な計算方法

イ ターゲットとすべき母集団サイズの算出

ターゲットとすべき母集団サイズの算出に当たっては、以下のとおり、枠母集団サイズを加減算することで計算する。

要件	具体例
「抽出時の階層」と「回答時の階層」等を比較して、大幅な乖離※が認められる場合は、実数で加減算し、それ以外の場合は、抽出率の逆数分の加減算を行う（8層を往来する場合は、全て、実数で加減算する。）。	【Case 1】 標本事業所の中に、抽出時は「1層」であったものの、回答時は「5層」であった事業所が1件存在するような場合 ・1層の母集団サイズ - 1件 ・5層の母集団サイズ + 1件 【Case 2】 標本事業所の中に、抽出時は「1層」であったものの、回答時は「2層」であった事業所が2件存在するような場合 ・1層の母集団サイズ - (2件 × 1層の抽出率の逆数) ・2層の母集団サイズ + (2件 × 1層の抽出率の逆数)

ロ 乗率の算出

※大幅な乖離とは、「抽出時の階層」と「回答時の階層」を比較し、2以上離れた階層となった場合を指す（2層→4層など）。


キャリブレーション（税務データの活用を除く。）に使用する乗率の算出に当たっては、以下のとおり、計算する。

$$\text{乗率（階層別）} = \frac{\text{算出したターゲットとすべき母集団サイズ（階層別）}}{\text{「STEP 1」と「STEP 2」で求めた乗率を掛け合せたものの回答時の階層ごとの総和}}$$

3 調査対象外事業所の推定方法について

「母集団名簿の整備」において、接触できなかった事業所に係る調査対象外事業所の取扱いについては、接触できなかった事業所に内在する調査対象外事業所の推定を実施する場合、「接触できた事業所に占める調査対象外事業所の割合」を使用した計算を行うことを検討しているが、当該割合は、実査の過程において把握した調査対象外件数の多寡に依存するため、各年分の状況に応じて、その値が、不安定となることが想定される。

そのため、調査対象外事業所の推定を行う場合は、より安定的な手法とすることを目的として、以下のとおり、移動平均の考え方を使用した平準化処理を実装することとしたい。

記号の定義	計算式（n年の計算を行う場合）
m_{ghk} : g国税局h層k階級の抽出事業所数	① n年、n-1年及びn-2年について、それぞれ、「 c_{ghk}^* 」を算出する。
b_{ghk} : g国税局h層k階級の回答事業所数	② 算出した「 c_{ghk}^* 」及び各年の「 m_{ghk} 」を使用して、全体に占める割合の3年平均値を算出する。
c_{ghk} : g国税局h層k階級の調査対象外事業所数	3年平均値の算出： $\frac{(c_{ghk(n)}^* + c_{ghk(n-1)}^* + c_{ghk(n-2)}^*)}{(m_{ghk(n)} + m_{ghk(n-1)} + m_{ghk(n-2)})} \cdots (1)$
d_{ghk} : g国税局h層k階級の未回収事業所数	③ 算出した当該平均値をn年の全体に乗じる計算を行い、平準化されたn年の「 c_{ghk}^* 」を算出する。
v_{ghk} : g国税局h層k階級の調査対象外の推定割合	最終的な計算： $c_{ghk(n)}^{(R)} = m_{ghk(n)} \times (1)$
c_{ghk}^* : g国税局h層k階級の調査対象外事業所数 及び推定した調査対象外事業数( の部分)	
<div style="text-align: center;"> m_{ghk} v_{ghk} </div> 