

資料 2

保存期間：5年  
(令和9事務年度末)  
令和4年10月7日

# 民間給与実態統計調査の 復元推計手法と進捗状況について

## 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

### 1 見直し内容の概要について（「第3回国税庁所管統計の整備に関する検討会」資料を基に作成）

各見直し項目の概要は以下のとおり。

#### 「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。
- ・また、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、当該サイズへ向けた復元推計となるよう計算を行う。

#### 「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割合に関す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。

#### 「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。

#### 「その他」

- ・復元推計手法の見直しを行い、参考計表として過年分調査に関する遡及修正を実施する場合は、過年分調査時に実施した給与所得者の復元方法の見直しや、集計内容についても併せて適用する。

（平成27年分）事業所規模の表章区分「10人未満」について、「1～4人」及び「5～9人」へ細分化。

（令和元年分）従来の復元方法（標本抽出率の逆数を乗じる方法）から、「事業所用の調査票」に記載された「実際の給与所得者数」を用いて復元する方法に変更。

（令和2年分）「乙欄適用者」を除いた統計表を集計・公表。

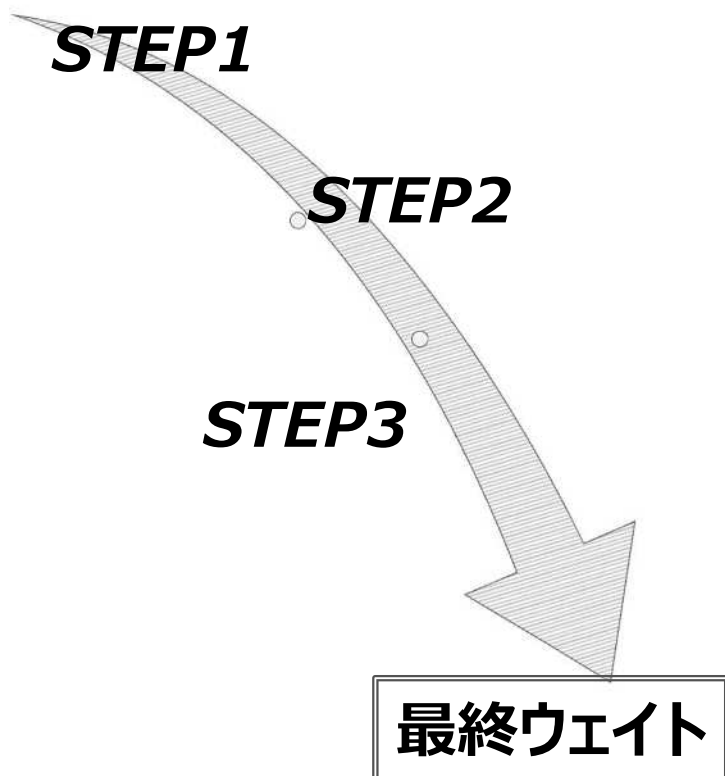
## 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

### 2 見直し後の復元推計方法の算出式について（「第1回及び第2回国税庁所管統計の整備に関する検討会」資料を基に作成）

#### （1）概要

復元推計方法の概要は以下のとおり。

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）を行う。



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

※ 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。

# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

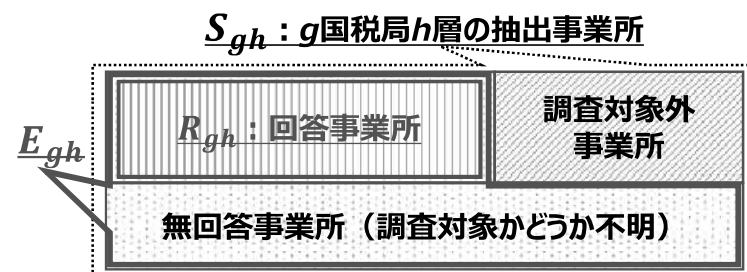
## (2) 記号の定義

### 母集団サイズ・抽出数・抽出率

- $M_{gh}$  :  $g$ 国税局 $h$ 層の枠母集団事業所数
- $S_{gh}$  :  $g$ 国税局 $h$ 層の抽出事業所
- $m_{gh}$  :  $g$ 国税局 $h$ 層の抽出事業所数
- $m_{ghk}$  :  $g$ 国税局 $h$ 層 $k$ 階級の抽出事業所数
- $e_{gh}$  :  $g$ 国税局 $h$ 層の抽出率 (=  $m_{gh}/M_{gh}$ )
- $X_{ghk}$  :  $g$ 国税局 $h$ 層 $k$ 階級のキャリブレーション（税務データの活用を除く）のための乗率
- $f_{ghj}$  :  $g$ 国税局 $h$ 層 $j$ 事業所における2千万円以下の給与所得者に係る抽出率（令和元年分調査以降※ :  $t_{ghj}/q_{ghj}$ ）
- $t_{ghj}$  :  $g$ 国税局 $h$ 層 $j$ 事業所における2千万円以下の抽出給与所得者数
- $q_{ghj}$  :  $g$ 国税局 $h$ 層 $j$ 事業所における2千万円以下の給与所得者数
- $w_{gh'}$  :  $g$ 国税局 $h'$ 規模の調整係数

### 調査結果

- $E_{gh}$  :  $g$ 国税局 $h$ 層の抽出事業所の内、回答事業所及び税務データで給与支給人員を補完した無回答事業所 ( $E_{gh} \subset S_{gh}$ )
- $R_{gh}$  :  $g$ 国税局 $h$ 層の回答事業所 ( $R_{gh} \subset E_{gh} \subset S_{gh}$ )
- $T_{ghj}$  :  $g$ 国税局 $h$ 層 $j$ 事業所における2千万円以下の抽出給与所得者
- $U_{ghj}$  :  $g$ 国税局 $h$ 層 $j$ 事業所における2千万円超の給与所得者
- $K_{gh}$  :  $g$ 国税局 $h$ 層の階級数 (1, 2又は3 ( $g$ と $h$ による))
- $b_{ghk}$  :  $g$ 国税局 $h$ 層 $k$ 階級の回答事業所数
- $b_{ghk}^*$  :  $g$ 国税局 $h$ 層 $k$ 階級の回答及び税務データ補完事業所数
- $c_{ghk}$  :  $g$ 国税局 $h$ 層 $k$ 階級の調査対象外事業所数
- $c_{ghk}^*$  :  $g$ 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
- $c_{ghk}^{(R)}$  :  $g$ 国税局 $h$ 層 $k$ 階級の推定後調査対象外事業所数
- $d_{ghk}$  :  $g$ 国税局 $h$ 層 $k$ 階級の未回収事業所数
- $v_{ghk}$  :  $g$ 国税局 $h$ 層 $k$ 階級の調査対象外の推定割合



# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

## (2) 記号の定義（続き）

### 調査結果（続き）

- $n_j$  :  $j$ 事業所の給与支給人員数
- $n_j^*$  :  $j$ 事業所の税務データ補完後給与支給人員数
- $z_j$  :  $j$ 事業所の給与総額等
- $y_{ji}$  :  $j$ 事業所*i*者の給与総額等（給与支給人員数を含む）
- $\theta_{jk}$  :  $j$ 事業所が*k*階級であれば「1」、それ以外であれば「0」をとる変数
- $\varphi_{jh'}$  :  $j$ 事業所が*h'*規模であれば「1」、それ以外であれば「0」をとる変数

### 推定量

- $\hat{N}_{gh'}$  :  $g$ 国税局*h'*規模の給与支給人員数
- $\hat{N}_{gh'}^*$  :  $g$ 国税局*h'*規模の税務データ補完後給与支給人員数
- $\hat{N}_{gh'}^{(R)}$  :  $g$ 国税局*h'*規模の比推定後給与支給人員数
- $\hat{Z}_{gh'}^{(R)}$  :  $g$ 国税局*h'*規模の比推定後給与総額等
- $\hat{Y}_{gh'}^{(R)}$  :  $g$ 国税局*h'*規模の比推定後給与総額等（比推定後給与支給人員数を含む）

$g$ 国税局			表章における規模 $h'$						
			1規模	2規模	3規模	4規模	5規模	6規模	7規模
抽出層 $h$	1層	1階級	-	-	-	-	-	-	-
		2階級	-	-	-	-	-	-	-
		3階級	-	-	-	-	-	-	-
	2層	1階級	-	-	-	-	-	-	-
		2階級	-	-	-	-	-	-	-
		3階級	-	-	-	-	-	-	-
	階級 $k$	3層	1階級	-	-	-	-	-	-
		4層	1階級	-	-	-	-	-	-
		5層	1階級	-	-	-	-	-	-
6層		1階級	-	-	-	-	-	-	
7層		1階級	-	-	-	-	-	-	
8層	1階級	-	-	-	-	-	-		

# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

## (3) 計算式

### ① 源泉徴収義務者用調査票の計算

$$\hat{N}_{gh'}^{(R)} = \sum_{h=1}^8 \sum_{k=1}^{K_{gh}} \sum_{j \in R_{gh}} \theta_{jk} \cdot \varphi_{jh'} \cdot n_j \cdot 1 / e_{gh} \cdot (m_{ghk} - c_{ghk}^{(R)}) / b_{ghk} \cdot w_{gh'} \cdot X_{ghk}$$

$$\hat{Z}_{gh'}^{(R)} = \sum_{h=1}^8 \sum_{k=1}^{K_{gh}} \sum_{j \in R_{gh}} \theta_{jk} \cdot \varphi_{jh'} \cdot z_j \cdot 1 / e_{gh} \cdot (m_{ghk} - c_{ghk}^{(R)}) / b_{ghk} \cdot w_{gh'} \cdot X_{ghk}$$

### ② 給与所得者用調査票の計算

$$\hat{Y}_{gh'}^{(R)} = \sum_{h=1}^8 \sum_{k=1}^{K_{gh}} \sum_{j \in R_{gh}} \theta_{jk} \cdot \varphi_{jh'} \cdot \hat{y}_j \cdot 1 / e_{gh} \cdot (m_{ghk} - c_{ghk}^{(R)}) / b_{ghk} \cdot w_{gh'} \cdot X_{ghk}$$

$$\hat{y}_j = \sum_{i \in T_{ghj}} y_{ji} \cdot 1 / f_{ghj} + \sum_{i \in U_{ghj}} y_{ji} \cdot 1$$

# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

## (4) その他

### ① 調査対象外事業所数の計算式

i 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。

※ 本来予定していた実査を実施できなかった場合は、調査対象外事業所の把握が不十分であることが想定されることから、その年分については、計算から除外する。


ii 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。

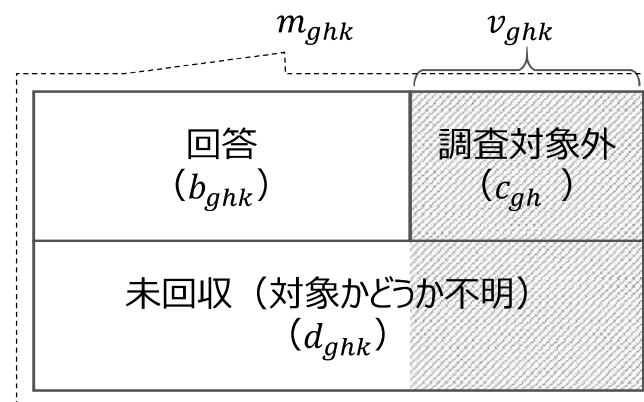
3年平均値の算出：

$$\frac{(c_{ghk(n)}^* + c_{ghk(n-1)}^* + c_{ghk(n-2)}^*) / (m_{ghk(n)} + m_{ghk(n-1)} + m_{ghk(n-2)})}{\dots} \quad \dots (イ)$$

iii 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。

$$\text{最終的な計算： } c_{ghk(n)}^{(R)} = m_{ghk(n)} \times (イ)$$

$m_{ghk}$	: g国税局h層k階級の抽出事業所数
$b_{ghk}$	: g国税局h層k階級の回答事業所数
$c_{ghk}$	: g国税局h層k階級の調査対象外事業所数
$d_{ghk}$	: g国税局h層k階級の未回収事業所数
$v_{ghk}$	: g国税局h層k階級の調査対象外の推定割合
$c_{ghk}^*$	: 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)



# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

## (4) その他（続き）

### ② 税務データを副次的情報として活用した比推定のための調整係数の計算式

i 全ての見直しを図った上で、給与支給人員を復元推計する。

$$\widehat{N}_{gh'} = \sum_{h=1}^8 \sum_{k=1}^{K_{gh}} \sum_{j \in R_{gh}} \theta_{jk} \cdot \varphi_{jh'} \cdot n_j \cdot 1 / e_{gh} \cdot (m_{ghk} - c_{ghk}^{(R)}) / b_{ghk} \cdot X_{ghk}$$

ii 税務データを活用し、無回答の事業所に係る給与支給人員を補完した上で給与支給人員を復元推計する。

$$\widehat{N}_{gh'}^* = \sum_{h=1}^8 \sum_{k=1}^{K_{gh}} \sum_{j \in E_{gh}} \theta_{jk} \cdot \varphi_{jh'} \cdot n_j^* \cdot 1 / e_{gh} \cdot (m_{ghk} - c_{ghk}^{(R)}) / b_{ghk}^* \cdot X_{ghk}$$

iii i 及び ii より、調整係数を算出する。

$$w_{gh'} = \widehat{N}_{gh'}^* / \widehat{N}_{gh'}$$



# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

## (4) その他（続き）

### ③ キャリブレーションに使用する乗率の計算式

#### i 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）。	<p><b>【Case 1】</b> 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</p> <ul style="list-style-type: none"> <li>・1層の母集団サイズ - 1件</li> <li>・5層の母集団サイズ + 1件</li> </ul> <p><b>【Case 2】</b> 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</p> <ul style="list-style-type: none"> <li>・1層の母集団サイズ - (2件 × 1層の抽出率の逆数)</li> <li>・2層の母集団サイズ + (2件 × 1層の抽出率の逆数)</li> </ul>

※ 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）。

#### ii 乗率の算出

キャリブレーション（税務データの活用を除く）に使用する乗率の計算式は、以下のとおり。

$$\text{乗率} (X_{ghk}) = \frac{\text{算出したターゲットとすべき母集団サイズ (階層別)}}{\text{「標本として選ばれる確率の逆数」と「回収になる確率の逆数」の乗率を掛け合せたものの回答時の階層ごとの総和}}$$

# 1 民間給与実態統計調査の見直し後の復元推計方法（計算式）まとめ

## (5) 標準誤差率の計算式

$$\widehat{V}(\hat{\theta}_H) = \left(1 - \frac{n_H}{N_H}\right) \frac{n_H - 1}{n_H} \sum_{i=1}^{n_H} (\hat{\theta}_{H(i)} - \hat{\theta}_H)^2$$

$$SE_H = \sqrt{\widehat{V}(\hat{\theta}_H)}$$

$$\text{標準誤差率} = SE_H / \hat{\theta}_H$$

### 母集団・有効回答事業所・推定値

$N_H$  H規模の推定した母集団サイズ

$n_H$  H規模の有効回答事業所数

$\hat{\theta}_H$  全ての有効回答事業所を使用して計算したH規模の推定値

$\hat{\theta}_{H(i)}$  有効回答事業所 (i) を除いて計算したH規模の推定値

### 計算値

$\widehat{V}(\hat{\theta}_H)$  H規模の推定値に係る分散※

$SE_H$  H規模の推定値に係る標準誤差

※ リサンプリング法（ジャックナイフ法）により算出。  
(標本事業所のうち、ひとつの標本事業所を除いて復元推計を行う方法を標本事業所数分繰り返し行い算出する。)

## 2 スケジュール

民間給与実態統計調査の復元推計手法の、今後の見直しスケジュールは以下のとおり。

項目	令和 4 年						令和 5 年									
	7月	8月	9月	10月	11月	12月	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月
マイルストーン	令和 3 年分公表 ▽						令和 4 年分公表 ▽									
復元推計手法の見直し	<div style="border: 1px solid black; padding: 10px; margin-bottom: 10px;"> <ul style="list-style-type: none"> <li>✓ 復元推計手法の見直しについては、<u>令和 4 年分調査から反映</u>。</li> <li>✓ その際、利用者ニーズに資するため、遡及可能な年分については、遡及修正を実施し、参考計表として、併せて公表することを予定。</li> </ul> </div> <div style="display: flex; justify-content: space-between;"> <div style="width: 80%;"> <p>●—————● 統計表を作成するシステムの改修作業</p> <p>●—————● 統計表の作成作業（過年分調査を含む）</p> <p>●- - - - -● 総務省への説明</p> </div> <div style="width: 15%; border: 2px solid black; text-align: center; vertical-align: middle;">           改善反映         </div> </div>															