

# 標本設計及び結果の推定方法

## 1 標本抽出方法

標本抽出方法は、都道府県を層、第1次抽出単位を平成27年国勢調査調査区とし、第2次抽出単位を世帯とする層化2段抽出法である。

第1次抽出では、47都道府県ごとに確率比例系統抽出により、全国で7,576調査区を抽出した。調査票Aに係る調査区と調査票Bに係る調査区は、次に示した方法により、独立に抽出を行っており、調査区数はそれぞれ7,152及び424である。

第2次抽出では、等確率系統抽出により、各調査区から12世帯を抽出した。

調査対象は、この方法により抽出された世帯にふだん住んでいる10歳以上の世帯員全員である。

### (1) 調査区の抽出（第1次抽出）

(ア) 標本調査区は、次のものを除く調査区の中から抽出した。

- a 山岳・森林・原野地帯等の調査区
- b 大きな工場・学校等のある調査区
- c 社会施設・大きな病院のある調査区
- d 刑務所・拘置所等のある調査区
- e 自衛隊地域の調査区
- f 駐留軍地域の調査区
- g 水面調査区

(イ) 調査区の抽出に当たっては、まず全国を47都道府県に区分した。各都道府県の標本調査区数は「都道府県別標本調査区数及び集計対象世帯数・人員一覧表」のとおりである。

(ウ) 都道府県ごとに、そこに含まれる調査区を次の基準により配列した（調査票Bに係る調査区の抽出の場合は①、④及び⑥による。）。

- ① 大都市圏に含まれるか否か
- ② 人口集中地区に含まれるか否か
- ③ 市町村の人口階級
- ④ 市区町村コード

⑤ 調査区の特長（国勢調査調査区番号の後置番号）

⑥ 国勢調査調査区番号

(エ) この配列を基に、都道府県ごとに全調査区の人口を累積し、累積した人口に対して確率比例系統抽出により、調査区を抽出した。

### (2) 調査世帯の抽出（第2次抽出）

調査区は、平均して約50世帯を含む地区である。

各標本調査区について、調査日の前に調査員が調査区内の全世帯を訪問し、調査区要図及び世帯一覧を作成した。この世帯一覧から、等確率系統抽出により12世帯を抽出した。

なお、やむを得ない理由により、調査を行うことができない世帯があった場合は、結果精度を確保するため、調査世帯を追加抽出した。

### (3) 「1日の生活時間」の調査日

調査は10月20日現在で行ったが、「1日の生活時間」に関しては、曜日ごとの結果を集計するため、標本調査区を無作為に次の8つのグループに分け、グループごとに10月16日から10月24日までの9日間のうち連続する2日間を調査日として指定した。

グループ番号	調査曜日	調査日
1	土日	10月16日、17日
2	土日	10月16日、17日
3	日月	10月17日、18日
4	火水	10月19日、20日
5	木金	10月21日、22日
6	金土	10月22日、23日
7	土日	10月23日、24日
8	土日	10月23日、24日

なお、曜日別の日数の配分については、各曜日の

生活時間配分の分散を考慮して設定した。

調査曜日	日数	延べ日数
平日（月火水木金）	各 1 日（金は 2 日）	6 日
土	5 日	5 日
日	5 日	5 日

## 2 結果の推定方法

調査結果のうち、「行動者数」及び 10 歳以上人口は、別途推計した地域（調査票 A に係る結果の場合は都道府県、調査票 B の場合は全国。以下同じ。）、男女、年齢階級別人口を基準人口とする比推定によった。

また、「延べ時間」も同様に算出し、「平均時間」は、この「延べ時間」を対応する人口（又は「行動者数」）の推定値で除して求めた。

「行動者率」は、推定値の百分比として算出した。

### (1) 推定値の算出

推定値の算出は、集計区分（調査票 A の「生活行動」、「生活時間・時間帯」、「平均時刻」及び調査票 B の「生活時間・時間帯」の別）、調査曜日別（「生活行動」以外）に次の手順により算出した。

#### ① 線型推定用乗率の算出

各地域について、標本調査区ごとに人口の線型推定用乗率を算出する。線型推定用乗率とは抽出単位の抽出率の逆数である。

ここで、第  $h$  地域、第  $i$  標本調査区の線型推定用乗率は、次の式で表される。

$$L_{hi} = \frac{1}{m_h} \cdot \frac{Q_h}{Q_{hi}} \cdot \frac{N_{hi}}{n_{hi}} \cdot r_{hi}$$

$L_{hi}$  : 第  $h$  地域、第  $i$  標本調査区の線型推定用乗率

$h$  : 地域

$i$  : 第  $h$  地域の標本調査区 ( $i=1,2,3,\dots,m_h$ )

$m_h$  : 第  $h$  地域の標本調査区数

$Q_h$  : 第  $h$  地域の国勢調査人口

$Q_{hi}$  : 第  $h$  地域、第  $i$  標本調査区の国勢調査人口

$N_{hi}$  : 第  $h$  地域、第  $i$  標本調査区の実査時の世帯

総数

$n_{hi}$  : 第  $h$  地域、第  $i$  標本調査区の集計対象世帯数

$r_{hi}$  : 第  $h$  地域、第  $i$  標本調査区の修正倍率  
(調査区の分割、合併があった場合の補正值)

#### ② 線型推定値の算出

地域、男女、年齢階級別 10 歳以上人口の線型推定値をそれぞれ算出する。すなわち、標本調査区ごとに、男女、年齢階級別調査人口にその調査区の線型推定用乗率を乗じ、これを地域内で合算することにより、各地域の男女、年齢階級別人口の線型推定値が得られる。

$$\hat{P}_h = \sum_i^{m_h} L_{hi} \cdot P_{hi}$$

$\hat{P}_h$  : 第  $h$  地域の人口の線型推定値

$P_{hi}$  : 第  $h$  地域、第  $i$  標本調査区の調査人口

#### ③ 比推定用乗率の算出

地域、男女、年齢階級別に、基準人口を線型推定値で除して比推定用乗率を算出する。

基準人口は、令和 2 年国勢調査結果、人口推計結果、平成 29 年患者調査結果などを基に総務省統計局で推計した令和 3 年 10 月 1 日現在の人口を用いた。

ここで、第  $h$  地域の比推定用乗率  $R_h$  は次の式で表わされる。

$$R_h = \frac{B_h}{\hat{P}_h}$$

$R_h$  : 第  $h$  地域の比推定用乗率

$B_h$  : 第  $h$  地域の（男女、年齢階級別）基準人口

#### ④ 推定値の算出

(ア) ある属性を持つ人口及び行動者数を推定するには、まず、各調査区について、男女、年齢階級別にその属性を持つ調査人口を求める。これに、①で求めた調査区ごとの線型推定用乗率を乗じ、さらに、③で求めた地域、男女、年齢階級別の比推定用乗率を乗じる。こうして得られた値を、地域内で合算することにより、その属性を持つその地域内の男女、年齢階級別の人口の推定値が得られる。これを必要に応じて、地域間、男女間、年齢階級間等で合算すれば目的

の属性を持つ人口の推定値が得られる。

なお、実際の集計に当たっては、線型推定用乗率に比推定用乗率を乗じたもの（集計用乗率という。）を各標本データに付与し、その集計用乗率を合算することにより各種属性別の推定値を算出している。

(イ)「延べ時間」の推定は、①から③の手順により得られた線型推定用乗率及び比推定用乗率を用いて行う。

すなわち、(ア)において各調査区の調査人口を求める代わりに、各調査区の調査人口の行動時間の合計（延べ時間）を求める。それ以降の手順は(ア)と同じである。

## (2) 「平均時間」等の算出

上記の方法で推定された人口（又は「行動者数」）及び「延べ時間数」を用いて、「平均時間」、「平均時刻」、「行動者率」を次の式により算出した。

### ① 「平均時間」

$$\left( \begin{array}{l} \text{属性}X\text{を持つ人口} \\ \text{の}Y\text{曜日における} \\ \text{活動}Z\text{の平均時間} \end{array} \right) = \frac{\left( \begin{array}{l} \text{属性}X\text{を持つ人口の}Y\text{曜} \\ \text{日における活動}Z\text{の延} \\ \text{べ時間の推定値} \end{array} \right)}{\left( \begin{array}{l} \text{属性}X\text{を持つ人口(又は行動者} \\ \text{数)の}Y\text{曜日における推定値} \end{array} \right)}$$

なお、平日及び週全体の平均時間については、この式で算出する曜日別結果の平均として算出した。

$$\begin{aligned} \text{平日} &= (\text{月曜平均} + \dots + \text{金曜平均}) / 5 \\ \text{週全体} &= (\text{月曜平均} + \dots + \text{日曜平均}) / 7 \end{aligned}$$

ただし、ある曜日に当該属性を持つ客体が存在しない場合は以下のとおりとした。

$$\begin{aligned} \text{平日総平均} &= \\ & (\text{月曜総平均} + \dots + \text{金曜総平均}) * / \\ & \text{客体が存在する曜日数} \end{aligned}$$

\*:客体が存在する曜日のみ

$$\begin{aligned} \text{平日行動者平均} &= \\ & (\text{月曜行動者平均} + \dots + \text{金曜行動者平均}) * / \\ & \text{行動者が存在する曜日数} \end{aligned}$$

\*:行動者が存在する曜日のみ

$$\begin{aligned} \text{週全体の総平均} &= \\ & (\text{平日総平均} \times 5 + \text{土曜総平均} + \text{日曜総平均}) / 7 \end{aligned}$$

なお、平日、土曜及び日曜のうち、1つでも当該属性を持つ客体が存在しない場合は、算出せず「-」とした。

$$\begin{aligned} \text{週全体の行動者平均} &= \\ & (\text{月曜行動者平均} + \dots + \text{日曜行動者平均}) * / \\ & \text{行動者が存在する曜日数} \end{aligned}$$

\*:行動者が存在する曜日のみ

### ② 「平均時刻」

各行動の開始又は終了時刻を「1日目午前0時から経過時間数」とし、「平均時刻」は、その「延べ時間」の推定値を対応する「行動者数」の推定値で除して求めた。

なお、平均時刻算出は連続する2日間の調査票が揃っているデータを集計対象とし、集計の曜日はその1日目の曜日を使用した。また、曜日別は平日、土曜日、日曜日の3区分とした。

### ③ 「行動者率」

$$\left( \begin{array}{l} \text{属性}X\text{を持つ人口の} \\ \text{活動}Z\text{の行動者率} \end{array} \right) = \frac{\left( \begin{array}{l} \text{属性}X\text{を持つ活動}Z\text{の行} \\ \text{動者数の推定値} \end{array} \right)}{\left( \begin{array}{l} \text{属性}X\text{を持つ人口の推定値} \end{array} \right)}$$

## 3 推定値の標本誤差

標本誤差の算出は、副標本方式により行った。

副標本方式による標本誤差の計算は、第1段の標本抽出において独立で均等な大きさの数組の標本を抽出しておくのが本来であるが、この調査では、その近似として事後的に4組の副標本を設定し、この4組の副標本ごとに算出された推定値を用いて、次式により算出した。

$$\hat{\sigma} = \sqrt{\frac{1}{4(4-1)} \sum_{k=1}^4 (\hat{x}_k - \hat{x})^2}$$

- $\hat{\sigma}$  : 推定値の標準誤差  
 $k$  : 副標本の番号 ( $k=1, 2, 3, 4$ )  
 $\hat{x}_k$  : 第  $k$  副標本に基づく推定値  
 $\hat{x}$  : 全標本に基づく推定値

主な項目の標準誤差率は、「主な項目の標準誤差率」の各表のとおりである。

表中の「-」は、以下のいずれかに該当する箇所である。

- ・ 該当数値のない箇所
- ・ 誤差集計を行うために設定した4組の副標本のうち、1組しか該当数値のない箇所
- ・ 曜日別の該当数値が1つでも欠けている箇所