

付 1 標本設計及び結果の推定方法

1 標本抽出方法

この調査の標本抽出方法は、第1次抽出単位を平成12年国勢調査調査区（以下、「調査区」という。）とし、第2次抽出単位を住戸とする層化2段抽出法である。

第1次抽出では、全国で約2万9千調査区を抽出し、第2次抽出では、それらの調査区の中から住戸を抽出した。

調査対象は、抽出された調査区の中から選定した抽出単位に居住する約44万世帯のうち15歳以上の世帯員全員である。

(1) 標本調査区の抽出（第1次抽出）

- ア 標本調査区は、次のものを除く全国の調査区（約88万調査区）の中から抽出した。
 - a 山林・原野・耕地のある区域（後置番号が2の調査区）
 - b 刑務所・拘置所のある区域（同5の調査区）
 - c 自衛隊区域（同6の調査区）
 - d 駐留軍区域（同7の調査区）
 - e 水面調査区（同9の調査区）
- イ 各都道府県ごとに、そこに含まれる調査区を産業及び従業上の地位別人口等によって層に分類し、都道府県別の各層ごとにウエイト（＝調査区内の世帯数）つき（不等確率）系統抽出を行った。

(2) 調査住戸の抽出（第2次抽出）

各標本調査区ごとに、その調査区のすべての住戸を列記した「抽出単位名簿」から、所定の抽出起番号と抽出間隔を用いて、等確率系統抽出法により調査住戸を抽出した。

なお、以上の方法で抽出された地域区分別の標本調査区数は、表1のとおりである。

2 結果の推定方法

結果の推定は、別途作成されている地域区分、男女、年齢階級（14区分）、世帯の種類（単身世帯、非単身世帯の2区分）別人口を基準人口とする比推定によった。

(1) 地域区分、男女、年齢階級、世帯の種類別人口の線形推定値の算出

次に示した96地域区分のそれぞれについて、男女、年齢階級、世帯の種類別人口の線形推定値を算出した。

- a 神奈川県及び福岡県
政令指定都市とそれ以外の地域の6地域
- b 神奈川県及び福岡県以外の都道府県
都道府県庁所在都市とそれ以外の地域の90地域

すなわち、各標本調査区の男女、年齢階級、世帯の種類別調査人口に、その調査区の線形推定用乗率（後述(3)を参照）を乗じ、これを地域区分内で合算して、各地域区分の男女、年齢階級、世帯の種類別人口の線形推定値を算出した。

(2) 比推定用乗率の算出

地域区分、男女、年齢階級、世帯の種類別に、基準人口を、それぞれ、前項(1)で求めた線形推定値で除して比推定用乗率を求めた。基準人口は、平成12年国勢調査結果、人口動態統計、人口移動統計などを基に、総務省統計局で推計した平成14年10月1日現在の人口を用いた。

なお、基準人口には次の者は含まれていない。

- a 外国政府の外交使節団、領事機関の構成員（随員を含む）及び外国軍隊の軍人・軍属並びにこれらの家族
- b 自衛隊の営舎内又は艦船内の居住者
- c 刑務所・拘置所に収容されている者のうち刑の確定している者及び少年院・婦人補導院の在院者

(3) 比推定値の算出

ア 特定の属性を持つ人口を推定するために、まず各調査区について、男女、年齢階級、世帯の種類別にその属性を持つ調査人口を求めた。これにその調査区の属する層の線形推定用乗率を乗じ、さらに前項(2)で求めた地域区分、男女、年齢階級、世帯の種類別の比推定用乗率を乗じた。こうして得られた値を地域区分内で合算して、その属性を持つその地域区分内の男女、年齢階級、世帯の種類別人口の比推定値を求めた。

これを必要に応じて、地域区分、男女、年齢階級、世帯の種類について合算した。

イ 地域区分、男女、年齢階級、世帯の種類別に特定の属性を持つ人口を推定する算式は、次のとおりである。

$$\begin{aligned}
\hat{X} &= \sum_{i=1}^L \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{W_i}{W_{ij}} f_{ij} X_{ij} \\
&\cdot \frac{P}{\sum_{i=1}^L \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{W_i}{W_{ij}} f_{ij} P_{ij}} \\
&= \sum_{i=1}^L F_i \sum_{j=1}^{m_i} \frac{f_{ij}}{W_{ij}} X_{ij} \\
&\cdot \frac{P}{\sum_{i=1}^L F_i \sum_{j=1}^{m_i} \frac{f_{ij}}{W_{ij}} P_{ij}} \\
&= R \sum_{i=1}^L \sum_{j=1}^{m_i} G_{ij} X_{ij}
\end{aligned}$$

ここで用いた記号の意味は次のとおりである。

\hat{X} : (地域区分, 男女, 年齢階級, 世帯の種類別) 当該属性を持つ人口の比推定値

i : 層の番号 ($i=1, 2, \dots, L$)

j : 第 i 層の第 j 標本調査区の番号
($j=1, 2, \dots, m_i$)

X_{ij} : 第 i 層の第 j 標本調査区内の当該属性を持つ (男女, 年齢階級, 世帯の種類別) 調査人口

W_{ij} : 第 i 層の第 j 標本調査区のウエイト (= 調査区内の世帯数)

f_{ij} : 第 i 層の第 j 標本調査区の世帯抽出率 (第 2 次抽出の抽出率) の逆数

m_i : 第 i 層の標本調査区数

W_i : 第 i 層に含まれるすべての調査区のウエイトの合計 (= $\sum_j W_{ij}$)

F_i : 第 i 層の線形推定用乗率 (= $\frac{W_i}{m_i}$)

G_{ij} : 第 i 層の第 j 標本調査区の線形推定用乗率
(= $F_i \cdot \frac{f_{ij}}{W_{ij}}$)

P_{ij} : 第 i 層の第 j 標本調査区内の (男女, 年齢階級, 世帯の種類別) 調査人口

P : (地域区分, 男女, 年齢階級, 世帯の種類別) 基準人口

R : (地域区分, 男女, 年齢階級, 世帯の種類別) 比推定用乗率 (= $\frac{P}{\sum_i \sum_j G_{ij} P_{ij}}$)

3 推定値の標本誤差

推定値の標本誤差は、全国についても、各地域区分についても、それぞれ 6 組の副標本ごとに独立に算出した推定値を用いて、次式により算出した。

$$\hat{s} = \sqrt{\frac{1}{6(6-1)} \sum_{k=1}^6 (\hat{X}_k - \hat{X})^2}$$

ここで用いた記号の意味は次のとおりである。

\hat{s} : 推定値の標本誤差

k : 副標本の番号 ($k=1, 2, \dots, 6$)

\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

このように計算された推定値に対する標準誤差の割合 (標準誤差率) は、表 2 ~ 4 のとおりである。