国民生活基礎調査の推計方法

（１）大規模調査年

ア　世帯票・健康票

各県（指定都市のある県については指定都市とそれ以外の地域のそれぞれを県とみなした。以下同じ。）の推計値$\hat{T}_k$は、世帯人員を補助変量とする比推定により、下記のように算定した。

$$\hat{T}_k = \frac{\sum_j X_{kj}}{\sum_j Y_{kj}} \cdot P_k$$

←**この意味については、Ｐ１５以降参照**

ただし、

$\hat{T}_k$ ： k県推計値

X_{kj} ： k県j調査地区のある属性を持つ世帯（員）数

Y_{kj} ： k県j調査地区内総世帯員数

P_k ： k県推計日本人人口（平成25年６月１日現在）

$\hat{T}_k$の分散の推計値は近似的に次式で与えられる。

$$V(\hat{T}_k) \fallingdotseq \hat{T}_k^2 \frac{(N_k - n_k)}{N_k n_k} \left\{ \frac{V_k(X)}{\overline{X}_k^2} - 2\frac{COV_k(X,Y)}{\overline{X}_k \cdot \overline{Y}_k} + \frac{V_k(Y)}{\overline{Y}_k^2} \right\}$$

ただし、

N_k ： k県国勢調査区数（後置番号１及び８）

n_k ： k県調査地区数

$$V_k(X) = \frac{1}{n_k - 1} \sum_j (X_{kj} - \overline{X}_k)^2$$

$$V_k(Y) = \frac{1}{n_k - 1} \sum_j (Y_{kj} - \overline{Y}_k)^2$$

$$COV_k(X,Y) = \frac{1}{n_k - 1} \sum_j (X_{kj} - \overline{X}_k)(Y_{kj} - \overline{Y}_k)$$

$$\overline{X}_k = \frac{\sum_j X_{kj}}{n_k}, \quad \overline{Y}_k = \frac{\sum_j Y_{kj}}{n_k}$$

$\hat{T}_k$の標準誤差の推計値は

$$\sqrt{V(\hat{T}_k)}$$

であり、標準誤差率の推計値は

$$\frac{\sqrt{V(\hat{T}_k)}}{\hat{T}_k}$$

で与えられる。

全国推計値$\hat{T}$は各県の推計値の合計とした。

即ち

$$\hat{T} = \sum_k \hat{T}_k$$

$\hat{T}$ ： 全国推計値

$\hat{T}$の分散の推計値は

$$V(\hat{T}) = \sum_k V(\hat{T}_k)$$

で求めた。

$\hat{T}$の標準誤差の推計値は

$$\sqrt{V(\hat{T})}$$

であり、標準誤差率の推計値は

$$\frac{\sqrt{V(\hat{T})}}{\hat{T}}$$

で与えられる。

大規模調査年においては都道府県別に表章するため、

① 調査年前年の 10 月 1 日現在の都道府県別日本人人口（総務省統計局人口推計より）を用いて、総務省の人口推計方法にもとづいて平成 25 年６月１日現在の都道府県別日本人人口を当部で推計する。

（理由：６月１日現在の都道府県別日本人人口が総務省統計局人口推計で作成されていないため。）

② 「①」の人口を調査結果から得られた都道府県・指定都市別世帯人員との比（拡大乗数）を求める。

③ 「②」の比（拡大乗数）を集落抽出により実施した調査結果から得られた世帯数及び世帯人員に乗ずる。

以上の計算によって都道府県・指定都市別の世帯数及び世帯人員を推計している。

例）平成 25 年大規模調査においての上記「②」
北海道（指定都市を除く）

北海道（指定都市を除く）の拡大乗数

$$\frac{\text{北海道（指定都市除く）平成 25 年 6 月 1 日現在推計人口（当部推計値）} \quad 3{,}490{,}765\text{ 人}}{\text{北海道（指定都市除く）平成 25 年国民生活基礎調査（世帯票）から得られた総世帯人員数} \quad 6{,}902\text{ 人}} = 505.761$$

（小数点第３位未満四捨五入）

イ　所得票・貯蓄票

推計値（ある属性を持つ世帯の平均所得、貯蓄等）$\hat{R}$は比推定により、下記のように算定した。

$$\hat{R}=\frac{\sum_k \left(\frac{N_k M_k}{n_k m_k} \sum_j X_{kj} \right)}{\sum_k \left(\frac{N_k M_k}{n_k m_k} \sum_j Y_{kj} \right)}$$

ただし、

$\hat{R}$ ： 推計値

N_k ： k 県国勢調査地区数（後置番号１）

n_k ： k 県世帯票調査地区数（後置番号１）

M_k ： k 県の n_k 個の調査地区から設定された単位区数

m_k ： k 県調査単位区数

X_{kj} ： k 県 j 単位地区のある属性をもつ世帯の総所得、貯蓄等

Y_{kj} ： k 県 j 単位地区のある属性をもつ世帯の総数

$\hat{R}$の分散の推計値は近似的に次式で与えられる。

$$V(\hat{R}) \fallingdotseq \hat{R}^2 \sum_k \left(\frac{L_k}{L} \right)^2 \left(\frac{1}{m_k} - \frac{1}{L_k} \right) \left\{ \frac{V_k(X)}{\bar{X}^2} - \frac{2COV_k(X,Y)}{\bar{X}\,\bar{Y}} + \frac{V_k(Y)}{\bar{Y}^2} \right\}$$

ここに、

$$L_k = \frac{N_k M_k}{n_k}, L = \sum_k L_k$$

ただし、

$$V_k(X)=\frac{1}{m_k-1}\sum_{i}(X_{ki}-\bar{X}_k)^2$$

$$V_k(Y)=\frac{1}{m_k-1}\sum_{i}(Y_{ki}-\bar{Y}_k)^2$$

$$COV_k(X,Y)=\frac{1}{m_k-1}\sum_{i}(X_{ki}-\bar{X}_k)(Y_{ki}-\bar{Y}_k)$$

$$\bar{X}_k=\frac{\sum_i X_{ki}}{m_k},\quad \bar{Y}_k=\frac{\sum_i Y_{ki}}{m_k},\quad \bar{X}=\sum_k\frac{L_k}{L}\bar{X}_k,\quad \bar{Y}=\sum_k\frac{L_k}{L}\bar{Y}_k$$

$\hat{R}$の標準誤差の推計値は

$$\sqrt{V(\hat{R})}$$

であり、標準誤差率の推計値は

$$\frac{\sqrt{V(\hat{R})}}{\hat{R}}$$

で与えられる。

大規模調査年においては、

① 都道府県・指定都市別の、国勢調査調査区数（後置番号１）と世帯票の実査地区数（後置番号１）の比及び世帯票実査地区から設定された単位区数と所得票の実査単位区数の比（拡大乗数）を求める。

② 「①」の比（拡大乗数）を集落抽出により実施した調査結果から得られた世帯数に乗ずる。

以上の計算によって１世帯当たりの平均所得金額等を推計している。

例）平成 25 年大規模調査においての上記「①」

北海道（指定都市を除く）

北海道（指定都市を除く）の拡大乗数

$$\frac{\text{北海道（指定都市除く）平成22年国勢調査調査区数（後置番号１）}\quad 29{,}277}{\text{北海道（指定都市除く）平成25年世帯票実査地区数（後置番号１）}\quad 75}\times\frac{\text{北海道（指定都市除く）平成25年世帯票実査地区から設定された単位区数}\quad 161}{\text{北海道（指定都市除く）平成25年所得票実査単位区数}\quad 41}=1532.877073$$

（２）簡易調査年

ア　世帯票

全国推計値 $\hat{Z}$ は、世帯人員を補助変数とする比推定により、下記のように算定した。

$\hat{Z}$ ：　ある属性を持つ世帯数（又は世帯員数）の全国推計値

X_{ij} ：　第 i 層の第 j 標本地区の当該属性をもつ世帯数（又は世帯員数）

Y_{ij} ：　第 i 層の第 j 標本地区の世帯員総数

N_i ：　第 i 層に含まれる国勢調査地区数（後置番号１及び８）

N ：　国勢調査地区数（後置番号１及び８）

n_i ：　第 i 層の標本地区数

n ：　標本地区総数

P ：　推計日本人人口（平成26年６月１日現在　125,480,777人総務省統計局「人口推計月報

とすると、全国推計値 $\hat{Z}$ は、

$$\hat{Z}=\frac{\sum_i \frac{N_i}{n_i}\sum_{j=1}^{n_i} X_{ij}}{\sum_i \frac{N_i}{n_i}\sum_{j=1}^{n_i} Y_{ij}}\cdot P \fallingdotseq \frac{\sum_i\sum_j X_{ij}}{\sum_i\sum_j Y_{ij}}\cdot P$$

で与えられる。

$\hat{Z}$ の分散の推計値は、近似的に、

$$V(\hat{Z}) \fallingdotseq \hat{Z}^2\frac{N-n}{Nn}\left\{\frac{V(X)}{\bar{X}^2}-2\frac{Cov(X,Y)}{\bar{X}\bar{Y}}+\frac{V(Y)}{\bar{Y}^2}\right\}$$

で与えられる。ここに、

$\bar{X}$, $\bar{Y}$ は、

$$\bar{X}=\frac{1}{n}\sum_i\sum_j X_{ij},\quad \bar{Y}=\frac{1}{n}\sum_i\sum_j Y_{ij}$$

$V(X)$、$V(Y)$、$Cov(X,Y)$ は、X, Y の分散及び共分散である。

$$V(X)=\frac{1}{n-1}\sum_i\sum_j (X_{ij}-\bar{X})^2,\quad V(Y)=\frac{1}{n-1}\sum_i\sum_j (Y_{ij}-\bar{Y})^2$$

$$Cov(X,Y)=\frac{1}{n-1}\sum_i\sum_j (X_{ij}-\bar{X})(Y_{ij}-\bar{Y})$$

$\hat{Z}$ の標準誤差の推計値は、

$$\sqrt{V(\hat{Z})}$$

であり、標準誤差率の推計値は

$$\frac{\sqrt{V(\hat{Z})}}{\hat{Z}}$$

で与えられる。

簡易調査年においては全国値で表章するため、

① 調査年の６月１日現在の日本人人口（総務省統計局人口推計より）と、調査結果から得られた世帯人員との比（拡大乗数）を求める。

② 「①」の比（拡大乗数）を集落抽出により実施した調査結果から得られた世帯数及び世帯人員に乗ずる。

以上の計算によって全国の世帯数及び世帯人員を推計している。

例）平成 26 年簡易調査においての上記「①」

拡大乗数

$$\frac{\text{平成 26 年 6 月 1 日現在推計人口（確定値）} \quad 125,480,777 \text{ 人}}{\text{平成 26 年国民生活基礎調査（世帯票）から得られた総世帯人員} \quad 116,456 \text{ 人}} = 1,077.495$$

（小数点第３位未満四捨五入）

イ　所得票

推計値（ある属性を持つ世帯の平均所得）$\hat{R}$は比推定により、下記のように算定した。

$\hat{R}$ ： 推計値

N_i ： 第 i 層に含まれる国勢調査地区数（後置番号１）

N ： 国勢調査地区数（後置番号１）

n_i ： 第 i 層の世帯票調査地区数（後置番号１）

n ： 世帯票調査地区数（後置番号１）

M_i ： 第 i 層のn_i個の調査地区から設定された単位区数

M ： n個の調査地区から設定された単位区

m_i ： 第 i 層の調査単位区数

m ： 調査単位区数

X_{ij} ： 第 i 層の j 単位地区のある属性をもつ世帯の総所得

Y_{ij} ： 第 i 層の j 単位地区のある属性をもつ世帯の総数

とすると、推計値$\hat{R}$は、

$$\hat{R}=\frac{\sum_i \frac{N_i M_i}{n_i m_i}\sum_j X_{ij}}{\sum_i \frac{N_i M_i}{n_i m_i}\sum_j Y_{ij}} \fallingdotseq \frac{\sum_i\sum_j X_{ij}}{\sum_i\sum_j Y_{ij}}$$

で与えられる。

$\hat{R}$の分散の推計値は、近似的に、

$$\mathrm{V}(\hat{R}) \fallingdotseq \hat{R}^2 \frac{L-m}{Lm}\left\{\frac{\mathrm{V}(X)}{\overline{X}^2} - 2\frac{\mathrm{Cov}(X, Y)}{\overline{X}\,\overline{Y}} + \frac{\mathrm{V}(Y)}{\overline{Y}^2}\right\}$$

で与えられる。ここに、

$$L=\frac{NM}{n}$$

$$\bar{X} = \frac{1}{m} \sum_i \sum_j X_{ij}, \quad \bar{Y} = \frac{1}{m} \sum_i \sum_j Y_{ij}$$

$$\mathrm{V}(X) = \frac{1}{m-1} \sum_i \sum_j (X_{ij} - \bar{X})^2$$

$$\mathrm{V}(Y) = \frac{1}{m-1} \sum_i \sum_j (Y_{ij} - \bar{Y})^2$$

$$\mathrm{Cov}(X, Y) = \frac{1}{m-1} \sum_i \sum_j (X_{ij} - \bar{X})(Y_{ij} - \bar{Y})$$

$\hat{R}$の標準誤差の推計値は

$$\sqrt{\mathrm{V}(\hat{R})}$$

であり、標準誤差率の推計値は

$$\frac{\sqrt{\mathrm{V}(\hat{R})}}{\hat{R}}$$

で与えられる。

簡易調査年においては、拡大乗数は求めていない。

$\boxed{\hat{T}_k = \frac{\sum_j X_{kj}}{\sum_j Y_{kj}} \times \mathrm{P}_k}$ の意味について

「回答者１人あたり世帯数」

ある県の世帯数（推定値）　＝ $\dfrac{\text{調査した世帯数}}{\text{調査した世帯の世帯員数}}$ 　×　ある県の人口

例①　人口 100 万人のある県で、1000 世帯（世帯員数：2500 人）調査。

「回答者１人あたり世帯数」は $\frac{1000\ \text{世帯}}{2500\ \text{人}} = 0.4$（世帯／人）。

その県の世帯数（推定）は

0.4 ×　100 万人　＝　40 万世帯

「拡大乗数」

式の形をかえて

ある県の世帯数（推定値）　＝　調査した世帯の数　× $\dfrac{\text{ある県の人口}}{\text{調査した世帯の世帯員数}}$

例②　拡大乗数は $\frac{1{,}000{,}000\ \text{人}}{2500\ \text{人}} = 400$

拡大乗数を用いると、例えば、

ある県の単独世帯の数　＝ 調査した単独世帯の数　×　拡大乗数

ある県の単独世帯以外の世帯の数＝ 調査した単独世帯以外の世帯の数 × 拡大乗数

と「属性ごとの」世帯数を計算できる。

例③　調査した 1000 世帯のうち

単独世帯　　　　　：500 世帯（世帯員数：　500 人）

単独世帯以外の世帯：500 世帯（世帯員数：2000 人）

とすると、

単独世帯の世帯数　　＝　500 世帯 × 拡大乗数 400　＝　20 万世帯

単独世帯以外の世帯数　＝　500 世帯 × 拡大乗数 400　＝　20 万世帯

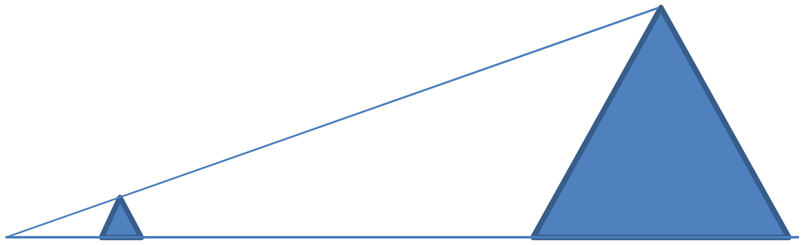
と推計。

単独世帯とそれ以外の世帯の構成割合

単独世帯の世帯数：単独世帯以外の世帯数＝１：１

(拡大乗数は無関係)

拡大乗数は、算数でいうところの単なる「比」にすぎない。



例④　現実には、すべての調査票が回収されるわけではない。単独世帯は５割しか回答が得られなかったとする。

単独世帯　　　　　　：250 世帯（世帯員数：　250 人）

単独世帯以外の世帯：500 世帯（世帯員数：2000 人）

すると、

「回答者１人あたり世帯数」　$\frac{750\text{世帯}}{2250\text{人}} = 0.333333\ldots$（世帯／人）。

例①に比べて「回答者１人あたり世帯数」は小さく見積られる。

そのため、この県の世帯数（推定）は

0.333333…　×　100 万人　≒　33 万世帯

と本来の 40 万世帯より、小さく推定される。

このように単独世帯の回収率が低いことによって、国民生活基礎調査の世帯数が小さく見積もられていると考えられる。

例⑤　また、拡大乗数 $\frac{1000000\text{人}}{2250\text{人}} \fallingdotseq 444.44$ なので

単独世帯の世帯数　　＝　250 世帯×拡大乗数 444.44　＝　111,111 世帯

単独世帯以外の世帯数＝　500 世帯×拡大乗数 444.44　＝　222,222 世帯

と推計。

単独世帯とそれ以外の世帯の構成割合

単独世帯の世帯数：単独世帯以外の世帯数＝１：２

（やはり拡大乗数は無関係）

このように世帯属性別回収率のばらつきによって、国民生活基礎調査の属性別世帯構成割合は、国勢調査のそれと違いが生じていると考えられる。