㉔ 集計法に基ずく偏倚の一例

遠　藤　健　児

§ 0　問　題

§ 1　標本抽出計画

§ 2　偏倚を除くための考慮

§ 3　抽出分散の推定値

§ 4　数　値　例

§ 附　録

1.　$\bar{x}'$ の積率

2.　抽出分散の推定

§ 0. 問題

調査単位を抽出する時の特殊事情のために，結果の集計に際して適当な修正を施す必要が生ずる場合のあることはよく知られている。　このような斟酌を行わないと偏倚を生じたり，誤差が殊更に大きくなったりする。　ここでは住宅調査に関して得られたデータ(*)によって，不完全なリストを使ったために入つくるかかる偏倚の大きさを数値的に評価してみた。

§ 1. 標本抽出計画

層別副次抽出法を用いることとして，第一次単位（市又は区）の抽出は確率比例方式によって，各層から一個宛を選ぶ。

即ち第一次単位の大きさをNとするとき，その推定値N_0に比例した確率を与えて抽出する。

抽出された第一次単位をΠとするとき，Πから第二次単位（調査単位）を抽出するのにはΠの完全なリストがないので，特殊なリストΠ'を用いて次のように標本抽出を行う。

Πの元とΠ'の元との間には一対一の対応はつけられないが，Πの各元はΠ'の上に少くとも一個の代表をもっているものとする。

かゝるΠ'から等確率にn個の単位を抽出して，対応するΠの元を求めてこれを調査単位と定める。　従って代表を2個或は2個以上持っている単位は重複して標本に選ばれることもあり得る。

予定標本数の割り当て，全体の抽出比rを予め定めておく。一つの層の大きさをΣN，その層に対応するリストに含まれる元で数えた大きさを$\Sigma N'$とすれば，この層の代表として選ばれた第一次単位からは，

(*) 昭25.12.1日の全国都市住宅調査（建設省住宅局）の調査表による。

$$n = (\Sigma N') \times r$$

個の調査単位を上述の方法で抽出する。 従って各層における標本数はリストの大きさN'による比例割当法で定めることになる。

§2. 偏倚を除くための考慮

標本として選ばれた第一次単位をΠとすると，Πに含まれる各単位は，それが何個の代表をリストΠ'に派出しているかによって分類されていると考えることができる。 ここで標本にとられた単位は，それを調査することによってどの類に属するかを確かめることができるものとする。 Πの元がもつ代表の個数をνとすれば，$\nu = 1, 2, 3, ---$ の各場合があって，ν個の代表をもつ元の集合をΠ_ν；その大きさをN_νとすれば，Πは$\Pi_1, \Pi_2, \Pi_3,$ ----- なる部分集団から成っていて，その大きさは

$$N = N_1 + N_2 + N_3 + ----$$

である。 一方Π'として示された集団は

$$\Pi_1 ; \Pi_2, \Pi_2 ; \Pi_3, \Pi_3, \Pi_3 ; -----$$

といった $1+2+3+$ ----- 個の部分集団の合併と考えられて，その大きさをN'とすれば

$$N' = N_1 + 2N_2 + 3N_3 + -----$$

である。

Π_νに属する単位が示す測度を

$$X_{\nu 1},\ X_{\nu 2},\ ----,\ X_{\nu N_\nu}$$

その合計並びに平均を夫々

$$X_\nu = \sum_j X_{\nu j}, \qquad \overline{X}_\nu = \sum_j X_{\nu j} / N_\nu$$

とおく。　Π 全体の平均 $\overline{X}$ は

$$\overline{X} = \sum_\nu N_\nu \overline{X}_\nu / N$$

である。

いま Π' によって大きさ n の標本を単純無作為に抽出したものとする。　これらの n 個はどの Π_ν に属するかによって分類することができる。　Π_ν に属するものが n_ν 個あったものとして，それらの測定値を

(1)　　$x_{\nu 1},\ x_{\nu 2},\ \text{-----},\ x_{\nu n_\nu}$　（$\nu = 1, 2, 3, \text{----}$）

で表わし，この値の合計並びに相加平均を夫々

(2)　　$x_{(\nu)} = \sum_{j=1}^{n_\nu} x_{\nu j}, \qquad \overline{x}_\nu = x_{(\nu)} / n_\nu$

とおく。

こゝで標本全体に亘る通常の算術平均

(3)　　$\overline{x} = \sum_\nu x_{(\nu)} / n$

をとると，これは Π' の平均，即ち

$$\overline{X}' = \sum_\nu \nu N_\nu \overline{X}_\nu / N'$$

の推定値であって，Π の平均 $\overline{X}$ の推定値としては

$$\overline{X}' - \overline{X}$$

という偏倚をもつことになる。　また Π が或る確率 P を以て第一次単位の標本として選ばれる場合にも，P の与え方に依る偏倚

の他にこの種の偏倚が入ってくる。

この種の偏倚を除くには，

$$(4)\qquad \bar{x}_0' = \frac{N'}{N}\cdot\frac{1}{n}\sum_{\nu}\frac{x_{(\nu)}}{\nu}$$

の形の推定値をとればよい。 しかし一般に云って N_ν，従って N' は未知であるから，乗数因子 $(Nn/N')^{-1}$ の代りにその推定値

$$\sum_{\nu}\frac{n_\nu}{\nu}$$

をとつた。

$$(5)\qquad \bar{x}' = \sum_{\nu}\frac{x_{(\nu)}}{\nu} \Big/ \sum_{\nu}\frac{n_\nu}{\nu}$$

を $\bar{X}$ の推定値として用いることができる。 この偏倚は n^{-1} の order であって，単純抽出の場合の $\bar{x}'$ の標準誤差は n^{-2} の order の項を省略すれば

$$(6)\qquad \sigma^2_{\bar{x}'} = \frac{1}{n}\cdot\frac{N'}{N}\cdot\frac{1}{N}\sum_{\nu}\frac{1}{\nu}\sum_{j}(X_{\nu j}-\bar{X})^2$$

$$= \frac{\sigma^2}{n}+\frac{1}{n}\left\{\frac{1}{N}\sum_{\nu}\left(\frac{N'}{N_\nu}-1\right)\sum_{j}(X_{\nu j}-\bar{X})^2\right\}$$

によって与えられる（附録1参照）。この右辺の第一項は一対一のリストによった場合の抽出分散である。 第二項は正負いずれともなり得る。 例えば Π が Π_1 と Π_2 とだけから成っていて，Π_2 における変動が Π_1 に比してかなり大きいような場合には，この方法は Π_2 から標本のとられる可能性を大きくしているので，精度は却って向上する。

確率比例式に一個の第一次単位を選んで，副次抽出を行う場合には，$\bar{x}$ の分散と同じ order の項の他に，第一次単位間の級間分散に相当する項が加わったものが抽出誤差の大きさを与える(*)

また層別を行つてしかも N' によつて比例割当てを行つたときには，$n : N'$ の比は層ごとに異らないから全標本から求めた (5) の形の $\bar{x}'$ は不偏推定値であつて（附録1参照），その抽出分散は，第一次単位間の分散の他に，各層での $\bar{x}'$ の分散の層の大きさを荷重とする平均から成る項が加わる。この項は次の様になる。（附録1）

$$(6)' \quad {}_{2}\sigma^{2}_{\bar{x}'} = \left\{ \frac{(\Sigma N')}{(\Sigma N)(\Sigma n)} \right\}^2 \Sigma \left\{ \frac{n}{N'} \sum_{\nu} \frac{1}{\nu} \sum_{j} (X_{\nu j} - \bar{X})^2 \right\}$$

ここで Σ は層全部に亘る和を示す。

§ 3. 抽出分散の推定値

$\bar{x}'$ の抽出分散のうちで n^{-1} に比例して減少する部分は (6) またはこの種の量の或る荷重平均として (6)' によつて与えられる。

一つの一次単位からとられた大きさ n の観測値 (1) に基づいて (6) の量を推定するには，

$$(7) \begin{cases} s^2 = \sum_{\nu} \frac{1}{\nu^2} \sum_{j} (x_{\nu j} - \bar{x}'')^2 \Big/ \left(\Sigma \frac{n_\nu}{\nu}\right)^2 \\ \text{但し} \quad \bar{x}'' = \sum_{\nu} \left(\frac{x_{(\nu)}}{\nu^2}\right) \Big/ \sum_{\nu} \left(\frac{n_\nu}{\nu^2}\right) \end{cases}$$

をとればよい。s^2 の *bias* は n^{-2} の *order* であることが示される。（附録2. 参照）.

s^2 を計算するには，

(*). 第一次単位の大きさ N と確率を定めるためのその推定値 N_0 との比が，単位毎にかなり異なるときには $\frac{1}{\bar{x}}$ には或る程度の偏倚が入るが，ここでは N と N_0 との比が余り違わないとしてこの偏倚を無視する。（前頁）

$$(8)\quad \left(\sum_{\nu} \frac{n_\nu}{\nu}\right)^2 s^2 = \sum_{\nu}\left(\frac{1}{\nu^2}\sum_{j} x_{\nu j}^2\right) - \bar{x}''\left(\sum_{\nu} \frac{x_{(\nu)}}{\nu^2}\right)$$

なる関係式に依ればよい。　なお比率の推定の場合は，

$x_{(\nu)} = n_\nu p_\nu$ とおき

$$(9)\quad \sum_{\nu}\left(\frac{x_{(\nu)}}{\nu^2}\right)\Big/\sum_{\nu}\left(\frac{n_\nu}{\nu^2}\right) = \sum_{\nu}\left(\frac{n_\nu p_\nu}{\nu^2}\right)\Big/\sum_{\nu}\left(\frac{n_\nu}{\nu^2}\right) = p''$$

とすれば，(8)の関係式は

$$(10)\quad \left(\sum_{\nu} \frac{n_\nu}{\nu}\right)^2 s^2 = \left(\sum_{\nu} \frac{n_\nu}{\nu^2}\right) p''\cdot(1-p'')$$

となる。

層別した場合，(5)の $\bar{x}'$ の分散のうちで，第一次単位内部での変動に基ずく成分たる(6)の量を推定するには

$$(11)\quad \left\{\sum\left(\sum_{\nu} \frac{n_\nu}{\nu}\right)\right\}^2 s^2 = \sum\left\{\sum_{\nu} \frac{1}{\nu^2}\sum_{j}\left(x_{\nu j} - \bar{x}''\right)^2\right\}$$

によって定まる s^2 を推定値にとればよい（附録2）

ここで $\sum$ は層全体に亘る和を表わす。　また比率の推定の場合には，(11)の関係式は次のようになる。

$$(12)\quad \left\{\sum\left(\sum_{\nu} \frac{n_\nu}{\nu}\right)\right\}^2 s^2 = \sum\left\{\left(\sum_{\nu} \frac{n_\nu}{\nu^2}\right) p''\cdot(1-p'')\right\}$$

この式から s^2 を求める計算は，各層で求めた(8)または(10)の右辺の和を，$\sum_{\nu} n_\nu/\nu$ の和の平方で割るだけでよい。

§4. 数値例

関東地方の市及び東京都の区が第一次抽出単位であつて，第一表の如く関東地方の市（東京都の立川，八王子を含む）は四つの層に，また東京都の区及び武蔵野市は三つの層に分けられ，その各層から一つ宛の市または区が住宅数（の推定値，N_0）による確率比例方式で抽出された。 全体を通じて凡そ $1/500$ の標本住宅が抽出されるように計画されて，第一次単位からの標本住宅の抽出は，住宅の戸別のリストΠがないために，世帯別のリストΠ'に依って行われたのである。

住宅に関する種々の特性のうちで，住宅の所有関係について 1. 持家（p_1），2. 借家（p_2），3. 給与住宅（p_3），4. その他（p_4） の各比率（括弧内の記号で示す）の推定，並びに一戸当りの平均居住人員$\overline{X}$と，平均畳数$\overline{Y}$の推定に関して，集計過程と計算結果とを第2表に示す。 第2表で，［p'または$\overline{x}'$］の行の数値が対応する量の推定値である。 ［pまたは$\overline{x}$］の行の数値は通常の平均による推定値である。 従つて二つの量の間の差は，pまたは$\overline{x}$の持つ偏倚の推定値である。［$\sum_{\nu} i_\nu$］の行には各列のνを荷重とする和を，また［$\sum_{\nu} i_\nu/\nu^2$］の行には$1/\nu$を荷重とする和を示す。 ［$\sigma_{p'}$ 又は $\sigma_{\overline{x}'}$］の行の数値は，(8)または(10)で求めた標準誤差*を求めるには，各類内での観測値の平方和 $\sum_j x_{\nu j}^2$ 或は $\sum_j y_{\nu j}^2$ が必要であるが，この数値は第3表に示した，* の推定値すである，ここで$\overline{x}'$または$\overline{y}'$の標準誤差

第4表は関東地方の市，及び東京都（区と武蔵野市よりなる）の各群について，層ごとの推定値と全体に亘る推定値を示す。

第5表には通常の平均値による推定値（偏倚を含む），また第6表には偏倚の大きさと，p'または$\overline{x}'$の標準誤差の推定値（(8)〜(12)で求めたもの ― 括弧内）とを示す。

偏倚 $\bar{X}'-\bar{X}$ の大きさは $\bar{x}-\bar{x}'$ で推定されるが，この抽出分散は $\bar{x}'$ や $\bar{x}$ のそれと同じ order である。 従って第6表の二種の数値を比較することによって，集計法に基づく偏倚の大きさが重大であるかどうかを見ることができる。

なお第1表に引用した資料によって，$\bar{X}$に関して7つの層の第一次単位間の分散を出してみると，第7表最右列のようになって，n^{-1} の order の級内分散の数倍程度である。

比率の推定では，いくつかの項の％の和が100となるようになっているから，偏倚は或程度相殺されるので，合成した推定値の bias はさほど重大ではないように見える。 しかし定量的な測度に関しては，入り込む偏倚はかなり重大である。

この例では，単独の層についても，或はその合併についても，偏倚の大きさが，その標準誤差の数倍乃至10数倍に及んでいる場合が少くない。 この偏倚は n に無関係であるから，標本数 n が大きいほどこの偏倚は抽出誤差に比して重大となってくる。

最後に標本としてとられた市又は区について，用いたリスト Π' がもし完全（Πとの対応が $1:1$）であったと仮定したときの推定値 $\bar{x}'(=\bar{x})$ 及び $\bar{y}'(=\bar{y}$ となる）の抽出分散 σ^2/n の推定値を求めて，これを第7表に示した。 リストが不完全な場合の抽出分散（(8)で求まる s^2）と比較して大して違っていない
従って，この例については，不完全なリストを用いたけれども，精度の点では別に変わりがないことがわかる。

附記．住宅の調査表を借して下さった建設省住宅局企画課の方々と，調査表の分類集計や種々の計算を手伝って下さった藤原長司，飯沼健司の両君並びに浜田辰子嬢に深く感謝します。

§ 附　　録.

1. $\overline{x}'$ の 積 率

Π' によって Π からとられた n 個の標本に関する観測値を(1)の如く分類する。　各類での頻度（ $n_1, n_2, \cdots, n_\nu, \cdots$ ）の分布は，同時的確率

$$
(13)\quad \begin{cases} P(n_1, n_2, \cdots, n_\nu, \cdots) = \dfrac{n!}{n_1!\,n_2!\cdots n_\nu!\cdots} \cdot \dfrac{(N_1)^{[n_1]}(2N_2)^{[n_2]}\cdots(\nu N_\nu)^{[n_\nu]}\cdots}{N^{[n]}} \\ \text{但し}\quad x^{[r]} = x(x-1)\cdots(x-r+1) \end{cases}
$$

によって与えられるから，この分布における二次の中心積率を，$\sigma(n_\mu, n_\nu)$ で表わすと

$$
(14)\quad \begin{cases} E(n_\nu) = n\,\dfrac{\nu N_\nu}{N'} \\ \sigma(n_\mu, n_\nu) = \dfrac{N'-n}{N'-1}\cdot n \cdot \dfrac{\nu N_\nu}{N'}\left(\delta_{\mu\nu} - \dfrac{\mu N_\mu}{N'}\right) \end{cases}
$$

となる。　ここで $\delta_{\mu\nu}$ はクロネッカーの記号である。

次に（ $x_{(1)}, x_{(2)}, \cdots, x_{(\nu)}, \cdots$ ）の分布における中心積率を求める。　（ $n_1, n_2, \cdots, n_\nu, \cdots$ ）を固定した上での條件附分布に関しては

$$
(15)\quad E(x_{(\nu)} \mid n_\nu) = n_\nu \overline{X}_\nu
$$

$$
(16)\quad \begin{cases} \sigma(x_{(\mu)}, x_{(\nu)} \mid n_\mu, n_\nu) = \delta_{\mu\nu}\cdot \dfrac{N'-n}{N'-1}\cdot n \cdot \dfrac{\nu N_\nu}{N'}\,\sigma_\nu^2 \\ \text{但し}\quad \sigma_\nu^2 = \dfrac{1}{N_\nu}\sum_j (X_{\nu j} - \overline{X}_\nu)^2 \end{cases}
$$

であるから

$$(17)\quad E(x_{(\nu)}) = E(n_\nu \overline{X}_\nu) = n\,\frac{\nu N_\nu}{N'}\,\overline{X}_\nu$$

となることは容易にわかる。　更に

$$(18)\quad \sigma(x_{(\mu)}, x_{(\nu)})$$
$$= E\left\{\sigma(x_{(\mu)}, x_{(\nu)} \mid n_\mu, n_\nu)\right\} + \sigma\left\{E(x_{(\mu)} \mid n_\mu), E(x_{(\nu)} \mid n_\nu)\right\}$$

なる関係から，(15), (16) を使って更に (14) の才二式を参照すると

$$(19)\quad \sigma(x_{(\mu)}, x_{(\nu)}) = \delta_{\mu\nu}\frac{N'-n}{N'-1}\cdot n\,\frac{\nu N_\nu}{N'}\sigma_\nu^2 + \overline{X}_\mu \overline{X}_\nu\,\sigma(n_\mu, n_\nu)$$

となることがわかる.

更にまた

$$(20)\quad \sigma(x_{(\mu)}, n_\nu) = \overline{X}_\mu\,\sigma(n_\mu, n_\nu)$$

となることも同様にして示される。

こゝで $x_{(1)}, x_{(2)}, \cdots, x_{(\nu)}, \cdots$ から成る一次形式統計量

$$(21)\quad x = \sum_\nu \frac{1}{\nu} x_{(\nu)}$$

と，$n_1, n_2, \cdots, n_\nu, \cdots$ から成る一次形式統計量

$$(22)\quad t = \sum_\nu \frac{1}{\nu} n_\nu$$

を考える。　(5) の $\overline{x}'$ はこの二つの統計量の比となっているから，Taylor 展開を利用して初めの主要項をとると

$$(23)\quad \sigma_{\overline{x}'}^2 = \frac{1}{E(t)^4}\sum_\mu\sum_\nu \frac{1}{\mu\nu}\left\{E(t)^2\sigma(x_{(\mu)}, x_{(\nu)}) - 2E(t)E(x)\sigma(x_{(\mu)}, n_\nu) + E(x)^2\sigma(n_\mu, n_\nu)\right\}$$

となる。

この右辺に (14), (18), (20) を代入すれば,

$$(24)\quad \begin{cases} E(x) = \sum_{\nu} \frac{1}{\nu} \cdot n \frac{\nu N_\nu}{N'} \overline{X}_\nu = n \frac{N}{N'} \cdot \overline{X} \\ E(t) = \sum_{\nu} \frac{1}{\nu} n \frac{\nu N_\nu}{N'} = n \frac{N}{N'} \end{cases}$$

なることに注意して,

$$(25)\quad \sigma_{\bar{x}'}^2 = \frac{N'-n}{N'-1} \cdot \frac{1}{n} \left\{ \frac{N'}{N} \cdot \frac{1}{N} \sum_{\nu} \frac{1}{\nu} \sum_{j} (X_{\nu j} - \overline{X})^2 \right\}$$

$$= \frac{N'-n}{N'-1} \cdot \frac{1}{n} \left\{ \sigma^2 + \frac{1}{N} \sum_{\nu} \left(\frac{N'}{N_\nu} - 1 \right) \sum_{j} (X_{\nu j} - \overline{X})^2 \right\}$$

なる結果を導くことができる. (23)で省略された項は n^{-2} の order である。

また $\bar{x}' = x/t$ の bias のうちで, その主要項は

$$\left\{ \frac{\sigma_t^2}{E(t)^2} - \frac{Cov(x,t)}{E(x)E(t)} \right\} \overline{X}$$

であるが, t や x の変動係数は $n^{-1/2}$ の order だから, この量が n^{-1} の order の量であることがわかる。

層別した場合には, 各層から求めた (4) の $\bar{x}_0'$ が層の平均の不偏推定値であるから

$$(26)\quad \bar{x}_T = \Sigma \left\{ \left(\frac{N}{\Sigma N} \right) \bar{x}_0' \right\} = \Sigma \left\{ \left(\frac{N}{\Sigma N} \right) \frac{N'}{N} \frac{1}{n} \sum_{\nu} x_{(\nu)} / \nu \right\}$$

は全体の平均の不偏推定値である. こゝで Σ は層全体に亘る和を示す. N'による比例割当てを行っていて $n : N'$ が層ごとに一定であると見做せるならば, この $\bar{x}_T$ は

$$(27)\quad \bar{x}_T = \frac{(\Sigma N')}{(\Sigma N)(\Sigma n)} \Sigma \left\{ \sum_{\nu} x_{(\nu)} / \nu \right\}$$

とかくことができる。　ここで $E(\bar{x}_T) = \Sigma N\bar{X}/\Sigma N$ なのであるが，この関係を $X \equiv 1$ なる場合で考えると

$$(28) \qquad E\left\{\Sigma\left(\sum_{\nu} n_\nu/\nu\right)\right\} = \frac{(\Sigma N)(\Sigma n)}{(\Sigma N')}$$

となる。従って(27)の $\bar{x}_T$ の右辺の常数因子をその推定値でおきかえた

$$(29) \qquad \bar{x}_T' = \frac{\Sigma\{\sum_{\nu} x_{(\nu)}/\nu\}}{\Sigma\{\sum_{\nu} n_\nu/\nu\}}$$

が，全体の平均の不偏推定値と見做せる。　この bias は n^{-1} の order であることが上と全様にして示すことができる。

(26)の $\bar{x}_T$ は，Nn/N' をその推定値 $\sum_{\nu} n_\nu/\nu$ でおきかえると次のようになる。

$$(30) \qquad \bar{x}_T = \Sigma\left\{\left(\frac{N}{\Sigma N}\right)\bar{x}'\right\}$$

ここで $\bar{x}'$ は(5)による荷重平均である。　この $\bar{x}_T$ によって抽出分散を求めてみると，(6)によれば

$$(31) \quad \sigma^2_{\bar{x}_T} = \Sigma\left\{\left(\frac{N}{\Sigma N}\right)^2 \frac{1}{n}\cdot\frac{N'}{N^2}\sum_{\nu}\frac{1}{\nu}\Sigma(X_{\nu j}-\bar{X})^2\right\}$$

となることがわかる。　これは各層で単純抽出を行った場合の分散である。　確率 $P = N/\Sigma N$ を与えて1個の第一次単位を抽出して，そこから副次抽出を行う場合には，第一次単位間の変動に基ずく項

$$(32) \qquad \Sigma\left\{P(\bar{X}-\Sigma P\bar{X})^2\right\}$$

が更に加わる，

N' による比例割当てが行われているとして

$$N'/n = \Sigma N' / \Sigma n$$

なる関係を用いて（31）の右辺を書き直すと（6)′が得られる。

2. 抽出分散の推定

$$(33) \quad \bar{x}'' = \sum_{\nu} \frac{x_{(\nu)}}{\nu^2} \Big/ \sum_{\nu} \frac{n_\nu}{\nu^2}$$

とおくと，$\bar{x}''$ は n_ν/ν^2 を荷重とする $\bar{x}_\nu$ の平均である。

従って

$$(34) \quad \sum_{\nu} \frac{1}{\nu^2} \sum_{j} (x_{\nu j} - \bar{x}'')^2 = \sum_{\nu}{}' \frac{1}{\nu^2} \sum_{j} (x_{\nu j} - \bar{x}_\nu)^2 + \Sigma' \frac{n_\nu}{\nu^2} (\bar{x}_\nu - \bar{x}'')^2$$

と書きかえることができる。 ここで Σ' は $n_\nu \geqq 1$ なる項のみについて加えていることを特に示したものである。

（$n_1, n_2, ----, n_\nu, ---$）を固定した上での條件附平均を E' で表わすことにすると，先づ

$$E'\left\{ \sum_{j} (x_{\nu j} - \bar{x}_\nu)^2 \right\} = \frac{n_\nu - 1}{\nu N_\nu - 1} \sum_{j} \nu (X_{\nu j} - \bar{X}_\nu)^2$$

であるから，（34）の第一項の條件附平均は

$$(35) \quad \sum_{\nu}{}' \frac{N_\nu n_\nu - N_\nu}{\nu N_\nu - 1} \cdot \frac{1}{\nu} \sigma_\nu^2$$

となる。 ここで σ_ν^2 は（16）で与えられる Π_ν 内の分散である。

次に $E'(\bar{x}_\nu) = \bar{X}_\nu$ であって，また $\bar{x}_\nu$ の條件附の分散は

$$\sigma'^2_{\bar{x}_\nu} = \frac{\nu N_\nu - n_\nu}{\nu N_\nu - 1} \frac{1}{n_\nu} \sigma_\nu^2$$

である。 かゝる量 $\bar{x}_1, \bar{x}_2, ---, \bar{x}_\nu, ---$ の間での偏差の平方

和（ n_ν/ν^2 を荷重にとる）が（34）の第二項である。

この型の量の平均は，

(36) $$\bar{X}'' = \sum_\nu \left(\frac{n_\nu}{\nu^2}\right)\bar{X}_\nu \Big/ \sum_\nu \left(\frac{n_\nu}{\nu^2}\right)$$

とおけば

(37) $$\sum_\nu{}' \left(\frac{n_\nu}{\nu^2}\right)\frac{\nu N_\nu - n_\nu}{\nu N_\nu - 1}\frac{1}{n_\nu}\sigma_\nu^2 - \sum_\nu{}' \left(\frac{n_\nu}{\nu^2}\right)^2 \frac{\nu N_\nu - n_\nu}{\nu N_\nu - 1}\frac{1}{n_\nu}\sigma_\nu^2 \Big/ \sum_\nu \left(\frac{n_\nu}{\nu^2}\right) + \sum_\nu \left(\frac{n_\nu}{\nu^2}\right)(\bar{X}_\nu - \bar{X}'')^2$$

となることが示される。(*)

（35）と，（37）の第一項とを加えると $\sum_\nu (n_\nu/\nu^2)\sigma_\nu^2$ となって，この平均は

$$E\left(\sum_\nu \frac{n_\nu}{\nu^2}\sigma_\nu^2\right) = \frac{n}{N'}\sum_\nu \frac{N_\nu}{\nu}\sigma_\nu^2$$

となる。　（37）の第二項については，分子，分母ともにその変動係数は n^{-1} の order であるから，平均をとると n^0 の order となる。　また（37）の第三項は

(*)　$x_1, x_2, \cdots, x_n$ について，

$$E(x_i) = \mu_i, \quad \sigma_{x_i}^2 = \sigma_i^2$$

であるとき，荷重 $w_1, w_2, \cdots, w_n$ によって

$$\bar{x} = \sum w_i x_i \Big/ \sum w_i$$

$$S = \sum w_i (x_i - \bar{x})^2$$

を作る。　ここで

$$\mu = \sum w_i \mu_i \Big/ \sum w_i$$

とおくと

$$E(\bar{x}) = \mu$$

$$\sigma_{\bar{x}}^2 = \sum w_i \sigma_i^2 - \sum w_i^2 \sigma_i^2 \Big/ \sum w_i + \sum w_i (\mu_i - \mu)^2$$

である。

$$\sum_\nu \frac{n_\nu}{\nu^2}(\bar{X}_\nu-\bar{X}'')^2=\sum_\nu \frac{n_\nu}{\nu^2}(\bar{X}_\nu-\bar{X})^2-\left(\sum_\nu \frac{n_\nu}{\nu^2}\right)^{-1}\left\{\sum_\nu \frac{n_\nu}{\nu^2}(\bar{X}_\nu-\bar{X})\right\}^2$$

であって，この第一項の平均は

$$E\left(\sum_\nu \frac{n_\nu}{\nu^2}(\bar{X}_\nu-\bar{X})^2\right)=\frac{n}{N'}\sum_\nu \frac{N_\nu}{\nu}(\bar{X}_\nu-\bar{X})^2$$

第二項の平均は n^0 の order である。

従って（34）の平均は

$$(38)\quad \frac{n}{N'}\sum_\nu \frac{N_\nu}{\nu}\left\{\sigma_\nu^2+(\bar{X}_\nu-\bar{X})^2\right\}$$
$$=\left(\frac{N}{N'}n\right)^2\frac{N'}{N}\frac{1}{n}\left\{\frac{1}{N}\sum_\nu\frac{1}{\nu}\sum_j(X_{\nu j}-\bar{X})^2\right\}+O(1)$$

となる，　一方

$$E\left(\sum_\nu \frac{n_\nu}{\nu}\right)=\frac{N}{N'}n$$

だから

$$(7)\quad \begin{cases} s^2=\sum_\nu \frac{1}{\nu^2}\sum_j(x_{\nu j}-\bar{x}'')^2\Big/\left(\sum\frac{n_\nu}{\nu}\right)^2 \\ \text{但し}\quad \bar{x}''=\sum_\nu \frac{n_\nu}{\nu^2}\bar{x}_\nu\Big/\sum\frac{n_\nu}{\nu^2}\end{cases}$$

とおくと

$$E(s^2)=\frac{1}{n}\cdot\frac{N'}{N}\frac{1}{N}\sum_\nu\frac{1}{\nu}\sum_j(X_{\nu j}-\bar{X})^2+O\left(\frac{1}{n^2}\right)$$

である．　即ち（7）の s^2 をとれば（6）の $\sigma_{\bar{x}}^2$ の不偏推定値が得られる。

層別した場合，N' による比例割当てを行っていて，$N':n$ が層ごとに変わらないならば，（29）の $\bar{x}_T'$ の抽出分散のうちの第一次単位内の変動に基づく部分は（6）′の式で与えられることは 1. で示された．ところで（6）′の右辺の最初の因数の推定値としては

(25)によれば $\left[\sum\left\{\sum_{\nu} n_\nu/\nu\right\}\right]^{-2}$ をとることができ，また第二の項に対しては，(38)をみると(34)を推定値にとることができる。従って n^{-1} より高次の項を省略すると(11)によって定まる s^2 が(6)'の不偏推定値ということになる。

もし用いたリスト Π' が完全（Π との対応が $1:1$）ならば，Π' は Π_1 だけから成ることになって(5)の $\bar{x}'$ は通常の相加平均 $\bar{x}$ となる。また(25)の右辺は，$N_\nu = 0$ ($\nu = 2, 3, \cdots$)，$N = N'$ だから

$$\sigma_{\bar{x}}^2 = \frac{N-n}{N-1}\frac{1}{n}\cdot\frac{1}{N}\sum_j (X_j - \bar{X})^2$$

となる。Π の分散を σ^2 とおくとこの式は

$$\sigma_{\bar{x}}^2 = \frac{N-n}{N-1}\frac{\sigma^2}{n}$$

とかける。実際に不完全なリストを用いて得た標本値(1)によって，Π の分散 σ^2 を推定するには，この項の初めに述べたのと全く同様に考えると，

$$(39) \qquad s^2 = \sum_{\nu}\frac{1}{\nu}\sum_j (x_{\nu j} - \bar{x}')^2 \Big/ \sum_{\nu}\frac{n_\nu}{\nu}$$

を推定値にとればよいことが示される。

第1表　層別表

層番号	関東 1	関東 2	関東 3	関東 4	東京都 1	東京都 2	東京都 3
第一次抽出単位（○印標本）	千葉	○横須賀	甲府	水戸	台東	○大田	○世田谷
	川口	鎌倉	宇都宮	○土浦	港	足立	杉並
	浦和	藤沢	○桐生	日立	新宿	品川	板橋
	市川	小田原	前橋	木更津	豊島	葛飾	目黒
	大宮	平塚	高崎	館山	○文京	北	中野
	船橋	茅ヶ崎	八王子		中央	江戸川	練馬
	○川越	熱海	熊谷		渋谷	墨田	武蔵野
	松戸	伊東	佐野		千代田	荒川	
	立川		足利			江東	
			太田				
			伊勢崎				
			栃木				
世帯数 No	159.718	139933	176200	64606	301849	431750	330614
住宅数 No'	136.237	120018	154802	58935	226789	340900	238809
標本数 n	298	182	306	123	570	891	618
抽出比 n/No'	1/536	1/769	1/576	1/525	1/530	1/485	1/535

（No．No'は23.8.1　建設省住宅調査資料による）

第2.1表　関東第一層（川越市）に関する集計

ν \ i_ν/ν	標本数 n_ν/ν	持家 $n_\nu p_{1\nu}/\nu$	借家 $n_\nu p_{2\nu}/\nu$	給与住宅 $n_\nu p_{3\nu}/\nu$	その他 $n_4 p_{4\nu}/\nu$	居住人員 $x_{(\nu)}/\nu$	畳数 $y_{(\nu)}/\nu$
1	262	96	161	5	0	1300	4019
2	13.5	7	6.5	0	0	102	302
3	2	1.33	0.67	0	0	17.67	58
4	0.5	0.5	0	0	0	5	21.25
11	0.09	0.09	0	0	0	2.64	8
$\sum_\nu i_\nu/\nu$	278.09	104.92	168.17	5.00	0	1427.31	4408.25
p' 又は $\overline{x}'$		37.7%	60.5%	1.8%	0	5.13	15.85
$\sum_\nu i_\nu$	298	117	176	5	0	1606	4970
p 又は $\overline{x}$		39.3%	59.1%	1.7%	0	5.39	16.68
$\sum_\nu i_\nu/\nu^2$	269.56	100.8	164.47	5.00	0	1358.38	4195.37
p'' 又は $\overline{x}''$		0.3712	0.6101	0.0185	0	5.04	15.56
σ_p 又は $\sigma_{\overline{x}'}$		2.9%	2.9%	0.8%	0	0.12	0.48
偏倚		1.6%	−1.4%	−0.1%		0.26	0.83

第2.2表　関東第二層（横須賀市）に関する集計

ν \ i_ν/ν	n_ν/ν	$n_\nu p_{1\nu}/\nu$	$n_\nu p_{2\nu}/\nu$	$n_\nu p_{3\nu}/\nu$	$n_\nu p_{4\nu}/\nu$	$x_{(\nu)}/\nu$	$y_{(\nu)}/\nu$
1	155	72	63	19	1	742	2268
2	12	8	3.5	0.5	0	86	267.5
3	0.67	0.33	0.33	0	0	7.67	18.67
5	0.2	0.2	0	0	0	3.40	6.8
$\sum_\nu i_\nu/\nu$	167.87	80.53	66.83	19.50	1.00	839.93	2560.97
p'又は$\bar{x}'$		48.0%	39.8%	11.6%	0.6%	5.00	15.26
$\sum_\nu i_\nu$	182	90	71	20	1	954	2893
p又は$\bar{x}$		49.5%	39.0%	11.0%	0.5%	5.24	15.90
$\sum_\nu i_\nu/\nu^2$	161.26	76.15	65.46	14.25	1.00	788.26	2409.33
p''又は$\bar{x}''$		0.4722	0.4022	0.1194	0.0062	4.89	14.94
$\sigma_{p'}$又は$\sigma_{\bar{x}'}$		3.8%	3.7%	2.5%	0.6%	0.16	0.62
偏倚		1.5%	−0.8%	−0.6%	−0.1%	0.24	0.64

第2.3表　関東第三層（桐生市）に関する集計表

i_ν/ν \ ν	n_ν/ν	$n_\nu p_{1\nu}/\nu$	$n_\nu p_{2\nu}/\nu$	$n_\nu p_{3\nu}/\nu$	$n_\nu p_{4\nu}/\nu$	$x_{(\nu)}/\nu$	$y_{(\nu)}/\nu$
1	264	118	134	11	1	1418	4229
2	16.5	2.5	14	0	0	130.5	340
3	2	0.33	1.33	0.33	0	19	45
4	0.5	0	0.5	0	0	6.75	39.75
8	0.13	0	0.13	0	0	3	5
$\Sigma i_\nu/\nu$	283.13	120.83	149.96	11.33	1.00	1577.25	4558.75
p' 又は $\bar{x}'$		42.7%	53.0%	4.0%	0.4%	5.57	16.45
Σi_ν	306	124	169	12	1	1797	5243
p 又は $\bar{x}$		40.5%	55.2%	3.9%	0.3%	5.87	17.13
$\Sigma i_\nu/\nu^2$	273.07	119.36	141.59	11.11	1.00	1491.65	442457
p'' 又は $\bar{x}''$		0.4371	0.5185	0.0406	0.0037	5.46	16.20
$\sigma_{p'}$ 又は $\sigma_{\bar{x}'}$		2.9%	2.9%	1.2%	0.3%	0.14	0.58
偏　倚		−2.2%	2.2%	−0.1%	−0.1%	0.30	0.68

第2.4表　関東第四層（土浦市）に関する集計表

ν \ i_ν/ν	n_ν/ν	$n_\nu p_{1\nu}/\nu$	$n_\nu p_{2\nu}/\nu$	$n_\nu p_{3\nu}/\nu$	$n_\nu p_{4\nu}/\nu$	$x_{(\nu)}/\nu$	$y_{(\nu)}/\nu$
1	110	73	22	4	11	567	1884
2	6	1	4	0.5	0.5	40	138.5
6	0.17	0	0.17	0	0	2.67	6.83
$\sum_\nu i_\nu/\nu$	116.17	74.00	26.17	4.50	11.50	609.67	2019.33
p'又は$\bar{x}'$		63.7%	22.5%	3.9%	9.9%	5.25	17.38
$\sum_\nu i_\nu$	123	75	31	5	12	663	2182
p又は$\bar{x}$		61.0%	25.2%	4.1%	9.6%	5.39	17.74
$\sum_\nu i_\nu/\nu^2$	113.03	73.50	24.03	4.25	11.25	587.45	1949.39
p''又は$\bar{x}''$		0.6503	0.2126	0.0376	0.0995	5.20	17.25
$\sigma_{p'}$又は$\sigma_{\bar{x}'}$		4.4%	3.7%	1.7%	2.7%	0.22	0.89
偏倚		−2.7%	2.7%	0.2%	−0.1%	0.14	0.36

第2.5表　東京第一層（文京区）に関する集計表

ν \ i_ν/ν	n_ν/ν	$n_\nu p_{1\nu}/\nu$	$n_\nu p_{2\nu}/\nu$	$n_\nu p_{3\nu}/\nu$	$n_\nu p_{4\nu}/\nu$	$x_{(\nu)}/\nu$	$y_{(\nu)}/\nu$
1	39.6	267	104	20	5	1831	4728
2	59.5	28	29	2	0.5	407.5	1113.5
3	12.67	4	7.67	0.67	0.33	116.33	280
4	2.75	2	0.75	0	0	.37	74.75
5	1.2	0.4	0.6	0	0.2	15.4	30.4
$\Sigma i_\nu/\nu$	472.12	301.40	142.02	22.67	6.03	2401.23	6226.65
p'又は$\bar{x}'$		63.8%	30.1%	4.8%	1.3%	5.09	13.19
Σi_ν	570	345	191	26	8	3196	8246
p又は$\bar{x}'$		60.5%	33.5%	4.6%	1.4%	5.61	14.47
$\Sigma i_\nu/\nu^2$	430.90	282.91	121.37	21.22	5.40	2084.36	5402.85
p''又は$\bar{x}''$		0.6566	0.2817	0.0492	0.0125	4.84	12.54
$\sigma_{p'}$又は$\sigma_{\bar{x}'}$		2.1%	2.0%	1.0%	0.5%	0.10	0.37
偏　倚		-3.2%	3.4%	-0.2%	0.1%	0.52	1.28

第2.6表　関東第二層（大田區）に関する集計表

ν \ i_ν/ν	n_ν/ν	$n_\nu p_{1\nu}/\nu$	$n_\nu p_{2\nu}/\nu$	$n_\nu p_{3\nu}/\nu$	$n_\nu p_{4\nu}/\nu$	$x_{(\nu)}/\nu$	$y_{(\nu)}/\nu$
1	633	382	161	86	4	2969	8393
2	91	46	33	11	1	648.5	1813.5
3	16.33	6.33	8.67	1.33	0	167.67	385.67
4	3.5	1.75	1	0.75	0	41.5	103.25
5	1.6	0.4	1.2	0	0	21	49.8
6	0.83	0.33	0.17	0.33	0	15	31.33
$\Sigma i_\nu/\nu$	746.26	436.81	205.04	99.41	5.00	3862.67	10776.55
p' 又は $\sigma_{\bar{x}}'$		58.5%	27.5%	13.3%	0.7%	5.18	14.44
Σi_ν	891	504	264	117	6	5130	14027
p 又は $\sigma_{\bar{x}}$		56.6%	29.6%	13.1%	0.7%	5.76	15.74
$\Sigma i_\nu/\nu^2$	685.28	407.69	180.91	92.19	4.50	3366.22	9469.30
p'' 又は $\bar{x}''$		0.5949	0.2640	0.1345	0.0066	4.91	13.82
$\sigma_{p'}$ 又は $\sigma_{\bar{x}}'$		1.7%	1.5%	1.2%	0.3%	0.08	0.30
偏　倚		-1.9%	2.1%	-0.2%	0	0.58	1.30

第2.7表　関東第三層（世田谷区）に関する集計表

i_ν/ν \ ν	n_ν/ν	$n_\nu p_{1\nu}/\nu$	$n_\nu p_{2\nu}/\nu$	$n_\nu p_{3\nu}/\nu$	$n_\nu p_{4\nu}/\nu$	$x_{(\nu)}/\nu$	$y_{(\nu)}/\nu$
1	396	180	159	54	3	1892	6293
2	85	49	33.5	2	0.5	593.5	1946.5
3	11.67	6.67	4.33	0.67	0	106	281
4	2.75	1.75	1	0	0	38.25	99.75
5	0.8	0.4	0.4	0	0	10.8	31
6	0.33	0.17	0.17	0	0	7.5	13.5
$\Sigma i_\nu/\nu$	496.55	237.99	198.40	56.67	3.50	2648.05	8664.75
p' 又は $\bar{x}'$		47.9%	40.0%	11.4%	0.7%	5.33	17.45
Σi_ν	618	308	246	60	4	3649	11664
p 又は $\bar{x}$		49.8%	39.8%	9.7%	0.7%	5.90	18.87
$\Sigma i_\nu/\nu^2$	443.30	207.27	177.55	55.22	3.25	2237.05	7393.18
p'' 又は $\bar{x}''$		0.4676	0.4005	0.1246	0.0073	5.05	16.68
$\sigma_{p'}$ 又は $\sigma_{\bar{x}'}$		2.1%	2.1%	1.4%	0.4%	0.10	0.43
偏　倚		1.9%	-0.2%	-1.7%	0	0.57	1.42

第3.1表 —— 類ごとの平方和 $\sum x^2_{\nu j} = s^2_\nu$

	川越	横須賀	桐生	土浦	文京	大田	世田谷
s_1^2	7452	4208	9034	3561	10369	16712	10658
s_2^2	1722	1366	2519	574	6353	10687	9415
s_3^2	513	277	551		3681	5753	3136
s_4^2	208		425		1590	2154	2241
s_5^2		289			1055	1397	794
s_6^2				256		1634	1013
s_8^2			576				
s_{11}^2	841						

第3.2表 — 類ごとの平方和 $\sum_j y^2_{\nu j} = s^2_\nu$

	川越	横須賀	桐生	土浦	文京	大田	世田谷
s_1^2	78201	43048	92741	42668	80018	154371	137375
s_2^2	15614	13571	18406	6211	56535	92727	105467
s_3^2	6060	1568	3089		23454	31789	23985
s_4^2	4093		12665		9173	15061	15525
s_5^2		1156			4256	8493	6166
s_6^2				1681		7330	3393
s_8^2			1600				
s_{11}^2	7744						

第４表 —— 全体に亘る推定値（比率は％で示す）

関東 層番号	n	p_1'	p_2'	p_3'	p_4'	$\bar{x}'$	$\bar{y}'$
1	298	37.7	60.5	1.8	0	5.13	15.85
2	182	48.0	39.8	11.6	0.6	5.00	15.26
3	306	42.7	53.0	4.0	0.4	5.57	16.45
4	123	63.7	22.5	3.9	9.9	5.25	17.38
関東	909	45.0	48.6	4.8	1.6	5.27	16.15
東京 層番号	n	p_1'	p_2'	p_3'	p_4'	$\bar{x}'$	$\bar{y}'$
1	570	63.8	30.1	4.8	1.3	5.09	13.19
2	891	58.5	27.5	13.3	0.7	5.18	14.44
3	618	47.9	40.0	11.4	0.7	5.33	17.45
東京都	2079	56.9	31.8	10.4	0.8	5.20	14.97
関東地方	2988	53.0	37.4	8.6	1.1	5.22	15.36

第5表 —— 偏倚を含む推定値（比率は％で示す）

関東層番号	n	p_1	p_2	p_3	p_4	$\bar{x}$	$\bar{y}$
1	298	39.3	59.1	1.7	0	5.39	16.68
2	182	49.5	39.0	11.0	0.5	5.24	15.90
3	306	40.5	55.2	3.9	0.3	5.87	17.13
4	123	61.0	25.2	4.1	9.8	5.39	17.74
関東	909	44.7	49.2	4.6	1.5	5.52	16.82
東京層番号	n	p_1	p_2	p_3	p_4	$\bar{x}$	$\bar{y}$
1	570	60.5	33.5	4.6	1.4	5.61	14.47
2	891	56.6	29.6	13.1	0.7	5.76	15.74
3	618	49.8	39.8	9.7	0.7	5.90	18.87
東京都	2079	55.7	33.7	9.8	0.9	5.76	16.32
関東地方	2988	52.3	38.4	8.2	1.1	5.69	16.47

第6表 —— 偏倚の大きさとその標準誤差（括弧内 — 比率については%で示す）

関東 層番号	n	p_1	p_2	p_3	p_4	$\bar{x}$	$\bar{y}$
1	298	1.6 (2.9)	−1.4 (2.9)	−0.1 (0.8)		0.26 (0.12)	0.83 (0.48)
2	182	1.5 (3.8)	−0.8 (3.7)	−0.6 (2.5)	−0.1 (0.6)	0.24 (0.16)	0.64 (0.62)
3	306	−2.2 (2.9)	2.2 (2.9)	−0.1 (1.2)	−0.1 (0.3)	0.30 (0.14)	0.68 (0.58)
4	123	−2.7 (4.4)	2.7 (3.7)	0.2 (1.7)	−0.1 (2.7)	0.14 (0.22)	0.36 (0.89)
関東	909	−0.3 (1.7)	0.6 (1.6)	−0.2 (0.7)	−0.1 (0.4)	0.25 (0.08)	0.67 (0.30)
東京 層番号	n	p_1	p_2	p_3	p_4	$\bar{x}$	$\bar{y}$
1	570	−3.2 (2.1)	3.4 (2.0)	−0.2 (1.0)	0.1 (0.5)	0.52 (0.10)	1.28 (0.37)
2	891	−1.9 (1.7)	2.1 (1.5)	−0.2 (1.2)	0 (0.3)	0.58 (0.08)	1.30 (0.30)
3	618	1.9 (2.1)	−0.2 (2.1)	−1.7 (1.4)	0 (0.4)	0.57 (0.10)	1.42 (0.43)
東京都	2079	−1.2 (1.1)	1.9 (1.1)	−0.6 (0.7)	0.1 (0.2)	0.56 (0.05)	1.35 (0.21)
関東地方	2988	−0.7 (0.9)	1.1 (0.9)	−0.4 (0.5)	0 (0.2)	0.47 (0.04)	1.11 (0.17)

第7表 —— 完全なリストに依る場合との抽出分散の比較

第一次単位	$\overline{x}$		$\overline{y}$		標本数	層における第一次単位間の $\overline{x}$ の分散 *
	不完全	完全	不完全	完全	n	
川越	0.0144	0.0156	0.2302	0.2399	298	0.0144
横須賀	0.0262	0.0275	0.3784	0.3766	182	0.0747
桐生	0.0201	0.0216	0.3345	0.3429	306	0.0809
土浦	0.0488	0.0483	0.7889	0.7669	123	0.0425
文京	0.0109	0.0117	0.1335	0.1379	570	0.0693
大田	0.0067	0.0076	0.0929	0.0938	891	0.0495
世田谷	0.0092	0.0104	0.1801	0.1713	618	0.0519

註 * 23.8.1 建設省住宅調査資料から求めたもの