RESEARCH MEMOIR OF THE BUREAU OF STATISTICS

March 1970 No. 1

CONTENTS

BUREAU OF STATISTICS
OFFICE OF THE PRIME MINISTER, JAPAN

参考4－1

統計局研究彙報

1971年2月
第20号

総理府統計局総合研究係

個人企業経済調査の層別集計について

井出　　満

目　　次

はじめに

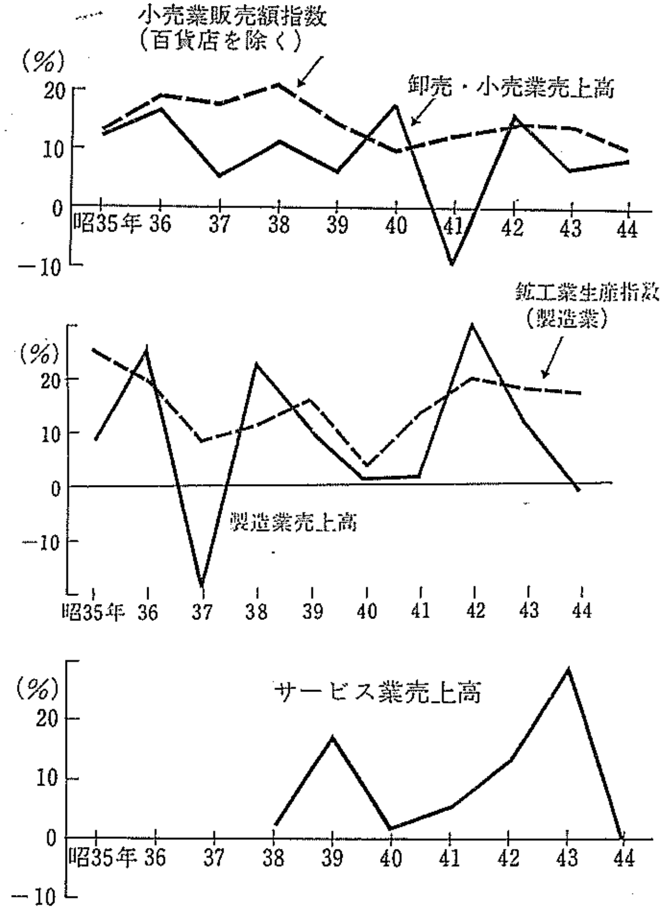
個人企業経済調査は，昭和22年に経済安定本部（現在の経済企画庁）が，国民所得の推計資料を得るために実施した「個人企業経済調査」から発展したもので，その後昭和27年に，総理府統計局に移管された。

このような歴史をもつ調査であったため，その初期の調査の第1の目的は，国民所得推計のための基礎資料を提供することであった。その観点からは，標本誤差が若干大きくてもあまり問題はなかった。しかし，その結果の利用が個人企業の経営の実態あるいはその時系列変化を明らかにするように発展してきたため，その大きな標本誤差が問題になった。そこで，昭和41年7月から標本数を拡大したが，まだ十分な標本数とはいえない。

ちなみに，現在の結果の標本誤差および時系列変化の他の統計との比較を示すと，第1図および第1表のとおりである。

このように，標本誤差が大きく，時系列分析等に耐える結果とはいいがたく標本数の一層の拡大が望まれるわけであるが，この調査は，調査員確保の困難

第１図　売上高の対前年上昇率



さ，記入拒否の多さ，記入内容の難かしさなど，実査上の問題が多々あり，また，集計工手間の問題もあって，標本数を容易に拡大するわけにはいかない。そこで，調査方法あるいは推定方法を工夫することによって，調査誤差および標本誤差を小さくすることを考える必要にせまられている。幸い，昭和46年度から集計方法が従来の手集計から機械集計に変わることになったので，若干複雑な推定方法も採用することが可能になった。そこで，本稿では，その一つの

第１表　主要項目の期別・全国平均の標準誤差率

（昭和43年第４四半期の結果より）単位：%

項　　目	F（製　造　業）	G（卸売・小売業）	L（サービス業）
売　上　高	9.3	8.2	8.2
仕　入　高	12.7	9.0	14.2
期首たな卸高	31.6	15.2	23.3
期末たな卸高	15.6	9.7	18.5
営　業　費	8.3	9.5	9.9
設　備　費	29.6	37.4	58.5

（注）標準誤差率を正確には計算していないので，調査結果より単位区間変動係数（C_b）および単位区内変動係数（C_w）を計算し，便宜的に次の算式を用いて求めた。

$$C(\bar{x})=\sqrt{\frac{C_b^2}{m}+\frac{C_w^2}{n}}$$

ここで，m は単位区数で178，n は調査客体数で製造業および卸売・小売業が1,000，サービス業が550とした。

推定方法であり，かなり実現性の高い層別集計（仮称）による推定方法について考察することにする。

Ⅰ　層別集計

1　考え方

層別集計の考え方は，標本抽出理論で，単純任意抽出法に比べ比例配分による層別単純任意抽出法の方が精度が良いという理論にヒントを得ている。実際に標本を抽出する際，調査単位を層別し，層ごとに調査単位を単純任意抽出するという層別単純任意抽出法を用いることができない場合，あるいはそのような抽出法を用いると，実査が困難な場合，集落抽出法とか多段抽出法など他の抽出法によらねばならない。その結果精度を特に考慮する配分をする場合は別として，層ごとの標本数は，母集団の層の大きさに比例しなくなるのが一般である。とくに標本数が少ない場合，その差が著しくなる。そこで，母集団の層の大きさを何らかの方法で知ることができれば，まず層ごとの標本平均を求め，その結果と母集団の層の大きさとを用いて全体の平均あるいは総和を推計する方法が考えられる。比例配分による層別単純任意抽

出法の効率の良さから，その結果が母集団の層の大きさを用いず単純に推計した結果より標本誤差が小さくなるであろうことは，類推できるであろう。

2　推定式と標準誤差

層別集計の推定式およびその標準誤差を，標本を単純任意抽出した場合について示すと，次のとおりである。

〔記　号〕

母集団の i 層 j 番目の調査単位の計量

$$X_{ij}\quad\begin{pmatrix} i=1,\ 2,\ \cdots\cdots,\ L \\ j=1,\ 2,\ \cdots\cdots,\ N_i \end{pmatrix}$$

標本の i 層 j 番目の調査単位の計量

$$x_{ij}\quad\begin{pmatrix} i=1,\ 2,\ \cdots\cdots,\ L \\ j=1,\ 2,\ \cdots\cdots,\ n_i \end{pmatrix}$$

母集団の大きさ

$$N=\sum_{i=1}^{L} N_i$$

標本の大きさ

$$n=\sum_{i=1}^{L} n_i$$

層の相対的大きさ

$$w_i=\frac{N_i}{N}$$

母平均

$$\mu=\frac{1}{N}\sum_{i=1}^{L}\sum_{j=1}^{Ni} X_{ij}$$

i 層の平均

$$\mu_i=\frac{1}{N_i}\sum_{j=1}^{Ni} X_{ij}$$

母分散

$$\sigma^2=\frac{1}{N}\sum_{i=1}^{L}\sum_{j=1}^{Ni}(X_{ij}-\mu)^2$$

i 層の分散

$$\sigma_i{}^2=\frac{1}{N_i}\sum_{j=1}^{Ni}(X_{ij}-\mu_i)^2$$

層間分散

$$\sigma_b{}^2=\frac{1}{N}\sum_{i=1}^{L} N_i\ (\mu_i-\mu)^2$$

層内分散

$$\sigma_w{}^2=\frac{1}{N}\sum_{i=1}^{L}\sum_{j=1}^{Ni}(X_{ij}-\mu_i)^2=\frac{1}{N}\sum_{i=1}^{L} N_i\sigma_i{}^2$$

〔推定式〕

$$\bar{x}'=\sum_{\substack{i=1\\(n_i\neq 0)}}^{L}\frac{N_i}{n_i}\sum_{j=1}^{ni} x_{ij}\Big/\sum_{\substack{i=1\\(n_i\neq 0)}}^{L} N_i\quad\left(\sum_{\substack{i=1\\(n_i\neq 0)}}^{L}\text{は }n_i=0\text{ になる層を除いた合計を意味する。}\right)$$

〔標準誤差〕

$$\sigma(\bar{x}')\doteqdot\sqrt{\frac{N-n}{N}}\cdot\frac{\sigma_w}{\sqrt{n}}$$

この推定式の期待値は，いずれかの層の標本数 n_i が0になる確率が小さければ，母平均 μ に近似する（注1）。また，標準誤差も n_i が0になる確率が小さければ層別単純任意抽出法と同じで，層間分散が関係しなくなる（注2）。そこで，単純任意抽出の標本誤差と比較するため，層別集計の単純集計に対する相対精度（η）を求めると，次のとおりとなり，層内分散を小さく層間分散を大きくするように層を作り層別集計すると，その効果が非常に上がることを示している（注3）。

〔層別集計の単純集計に対する相対精度〕

$$\eta\doteqdot 1+\frac{\sigma_b{}^2}{\sigma_w{}^2}$$

3　個人企業経済調査の集計への適用

層別集計を個人企業経済調査の集計に適用する場合を考えてみよう。個人企業経済調査にとって最も重要な調査事項および集計事項は，売上高である。したがって，前述したように，売上高の層間分散を大きく，層内分散を小さくするような指標を用いて層を作ればよいわけであるが，考えられるお

もな指標として，①従業者規模，②産業，③地方などがある。また，これらの指標により分類された母集団の大きさが必要であるが，これは，3年ごとにセンサスとして行なわれている事業所統計調査から得られる。

そこで，これらの指標を用いて層別集計をした場合，どのくらい効果が上がるか，昭和43年第4四半期の結果を用いて計算すると，次のとおりである。

第2表　売上高に関する層別集計の効果

（昭和43年第4四半期の結果より）　単位：千円

		全分散(σ^2)	層間分散 (σ_b^2)	層内分散 (σ_w^2)	$1+\frac{\sigma_b^2}{\sigma_w^2}$
F	従業者規模	$3,738^2$	$2,438^2$	$2,833^2$	1.74
	産業	$3,738^2$	616^2	$3,673^2$	1.03
	地方	$3,738^2$	514^2	$3,702^2$	1.02
	従業者規模×産業	$3,738^2$	$2,970^2$	$2,270^2$	2.71
G	従業者規模	$2,803^2$	$1,508^2$	$2,363^2$	1.41
	産業	$2,803^2$	858^2	$2,668^2$	1.10
	地方	$2,803^2$	178^2	$2,797^2$	1.00
	従業者規模×産業	$2,803^2$	$1,753^2$	$2,187^2$	1.64
L	従業者規模	733^2	426^2	597^2	1.51
	産業	733^2	258^2	685^2	1.14
	地方	733^2	47^2	731^2	1.00
	従業者規模×産業	733^2	520^2	517^2	2.01

（注）従業者規模
1人，2，3，4，5～9，10～19，20人～
産業
F，Gは中分類，Lは77洗たく・理容・浴場業については小分類，その他については中分類
地方
北海道・東北，関東，北陸，東海，近畿，中国・四国，九州

この結果からわかるように，個人企業の売上高は，従業者規模と非常に相関が高く，さらに産業を加味すると，層別集計の効果は非常に高まる。すなわち，従業者規模×産業の層別集計の単純集計に対する相対精度は，Fが2.71，Gが1.64，Lが2.01を示している。これは，たとえばFの場合，単純集計の標準誤差率が第1表で示したように9.3％であれば，層別集計の標準誤差率は，9.3％÷$\sqrt{2.71}$=5.5％というように，3.8ポイントも小さくなることを意味している。

Ⅱ　問題点

層別集計は，このように標本誤差を小さくすることができるという，非常な長所を持っているが，実際にこの集計方法を採用する場合，次のような偏りの問題を解決あるいは検討する必要がある。

1　$n_i=0$ の場合の偏り

層別集計による推定式は，いずれかの層の標本数 n_i が 0 になる確率が小さければ，偏りはあまり問題にならないが，その確率が大きいと不偏性が問題になってくる。とくに，層の数が多過ぎたり，層の作り方が悪いため，層の大きさが非常に小さいような場合，あるいは標本数があまり多くない場合 $n_i=0$ になる確率は大きい。そこで，まず $n_i=0$ となる確率について考えると，次のとおりとなる（注4）。

$$P_r(n_i=0)=\sum_{k(1)=1}^{L}(1-w_{k(1)})^n-\sum_{k(1)=1}^{L}\sum_{k(2)\neq k(1)}^{L}(1-w_{k(1)}-w_{k(2)})^n$$

$$+\cdots\cdots+(-1)^{L-2}\sum_{k(1)=1}^{L}\sum_{k(2)\neq k(1)}^{L}\cdots\cdots\sum_{k(L-1)\neq k(L-2)\neq\cdots\neq k(1)}^{L}(1-w_{k(1)}-w_{k(2)}-\cdots\cdots-w_{k(L-1)})^n$$

$$\fallingdotseq\sum_{k(1)=1}^{L}(1-w_{k(1)})^n$$

これに実際の値を与えて，$n_i=0$ となる確率を求めてみよう。たとえば

$w_i=\dfrac{1}{L}$

$L=2$ または 4

$n=10$ または100

の場合，$Pr(n_i=0)$ の値を計算すると，次のようになる。

$L=2$，$n=10$ のとき　$Pr(n_i=0)=1.953\times10^{-3}$

$n=100$ のとき　$Pr(n_i=0)=3.155\times10^{-30}$

$L=4,\ n=10$ のとき $P_r(n_i=0)=2.204\times10^{-1}$

$n=100$ のとき $P_r(n_i=0)=1.283\times10^{-12}$

したがって，層の大きさ w_i が大きければ大きいほど $P_r(n_i=0)$ は小さく，また，標本数 n が大きければ大きいほど $P_r(n_i=0)$ は小さくなることを示している。

層の大きさは，上記のようにすべて等しいことはないので，一番小さい層の大きさを w_1 と考え，$P_r(n_1=0)$ の値を求めてみよう。標本数としては，個人企業経済調査では，F と G が約1,000，L が550であるから，これら標本数の場合について計算してみると，次表のとおりである。

第3表　$P_r\ (n_1=0)$ の 値

w_1	$P_r(n_1=0)$	
	$n=1,000$	$n=550$
1/10,000	9.049×10^{-1}	9.465×10^{-1}
1/1,000	3.677×10^{-1}	5.768×10^{-1}
2/1,000	1.351×10^{-1}	3.325×10^{-1}
3/1,000	4.957×10^{-2}	1.916×10^{-1}
4/1,000	1.817×10^{-2}	1.102×10^{-1}
5/1,000	6.655×10^{-3}	6.349×10^{-2}
6/1,000	2.435×10^{-3}	3.652×10^{-2}
7/1,000	8.896×10^{-4}	2.099×10^{-2}
8/1,000	3.249×10^{-4}	1.206×10^{-2}
9/1,000	1.185×10^{-4}	6.927×10^{-3}
1/100	4.317×10^{-5}	3.975×10^{-3}
1/10	1.748×10^{-46}	6.814×10^{-26}

このように，個人企業経済調査の場合，層の大きさを F，G では 4/1,000 以上，L では 8/1,000 以上になるように作れば，偏りの問題はあまり気にする必要がないと考えてよい。

2　層の大きさを固定するための偏り

層の大きさが，その時その時分かっていれば問題ないが，一般に過去の結果しか分からないことが多い。実際，個人企業経済調査の場合も，層の大きさが分かるのは，3年ごとに行なわれる事業所統計調査の結果であって，毎月，毎期，毎年の層の大きさを把握することができない。そこで，3年間層の大きさを固定して層別集計すると，現実とかけ離れてくるため，その結果が偏りをもつことになる。これを算式で示せば，次のとおりである。

母集団の i 層の真の大きさを $\tilde{N}_i$ とすると，

$$\mu=\frac{1}{\tilde{N}}\sum_{i=1}^{L}\tilde{N}_i\,\mu_i=\sum_{i=1}^{L}\tilde{w}_i\,\mu_i$$

$$E(\bar{x}')=\frac{1}{N}\sum_{i=1}^{L}N_i\,\mu_i=\sum_{i=1}^{L}w_i\,\mu_i$$

$$Bias(\bar{x}')=E(\bar{x}')-\mu=\sum_{i=1}^{L}w_i\,\mu_i-\sum_{i=1}^{L}\tilde{w}_i\,\mu_i=\sum_{i=1}^{L}(w_i-\tilde{w}_i)\mu_i$$

ここで　$w_i=N_i/N,\ \tilde{w}_i=\tilde{N}_i/\tilde{N}$

実際にその偏りの大きさをみるため，従業者規模別に層を作った場合について，3年ごとに事業所統計調査の結果によって層の大きさを変更する際，従来どおりの層の大きさを用いて得られた推計値と，層の大きさを変更して推計した値とを対比してみると，次表のとおりである。

第4表　層の大きさの変更による売上高の違い（全国）

単位：千円

	F			G			L		
	①新	②旧	(②−①)/①	①新	②旧	(②−①)/①	①新	②旧	(②−①)/①
昭和33年平均	2,672	2,277	−14.8%	2,760	2,533	−8.2%	—	—	—%
36	4,245	4,020	− 5.3	3,629	3,568	−1.7	—	—	—
39	4,823	5,050	4.7	4,598	4,485	−2.5	1,314	1,309	−0.4
42	6,952	6,965	0.2	6,037	5,717	−5.3	1,620	1,558	−4.4

（注）「新」とは昭和33年が32年，36年が35年，39年が38年，42年が41年の事業所統計調査の結果を用いて推計したもので，「旧」とはそれぞれ29年，32年，35年，38年の結果を用いて推計したものである。

売上高の四半期別結果の標本誤差は，95%の信頼水準で，Fが18.6%，Gが16.4%，Lが16.4%となっており，上記の偏りに比べて非常に大きい。と

くに，時系列分析をする場合，第４表の偏りは３年間の累積であり、各期あるいは各年の対前期比，対前年比などの偏りの影響は，さらに小さくなるわけである。しかし，３年ごとに層の大きさを変更することは，第４表に示しただけの断層を生じさせるわけで，利用者にとって迷惑であることは間違いない。

おわりに

以上みてきたように，層別集計には，標本誤差を小さくするという利点と，偏りが生ずるという欠点とがあり，これをどのように評価するかにより，層別集計を採用するかどうか決定される。

しかし，いずれにしても，個人企業経済調査の標本誤差はあまりにも大きく，このような層別集計などの採用によりその誤差を小さくする必要がある。偏りについては，たとえば次のように解決すればよい。

まず，ある層に標本がえられない場合の偏りは，前述したように，層の大きさをＦ，Ｇでは 4/1,000 以上，Ｌでは 8/1,000 以上になるように，従業者規模と産業をクロスした層を作ればよい。その際，従業者規模の方が売上高との相関が高いので，売上高が近似している産業をまとめるような方法で層を所定の大きさ以上にするよう作ることにする。そして，各層に入る標本数 n_i と層の大きさ N_i とから，乗率 α_i を次のように定め，その乗率を用いて企業あたり平均値を推定すればよい。

$$\alpha_i=\frac{N_i}{n_i}\ \text{（4捨5入で整数とする）}$$

$$\bar{x}'=\frac{\sum_{i=1}^{L}\sum_{j=1}^{ni}\alpha_i x_{ij}}{\sum_{i=1}^{L} n_i\alpha_i}$$

ただし，n_i が０の場合，$\alpha_i=0$ として処理する。

次に，３年間層の大きさを固定するために生ずる偏りについては，前述したように，その偏りの大きさは，標本誤差にくらべそれほど大きくないので，３年ごとに層の大きさを変更するために生ずるギャップを善処すればよい。そこで，層の大きさを変更する年には，新旧の層の大きさを用いて重複集計し，そのギャップを明らかにするとともに，時系列の利用に供するため，おもな項目について，指数の形で表わし，層の大きさを変更する年には，新旧の層の大きさで集計した結果で，次のようにリンクすればよい。

$$\frac{\text{比較時の結果}}{\text{基準時の年平均の結果}}\times\text{リンク係数}$$

$$\text{リンク係数}=\frac{\text{旧の層の大きさによる年平均の結果}}{\text{新の層の大きさによる年平均の結果}}$$

（注１） 期待値を求める際，層ごとの標本数が定まったとしての期待値（E_2）と標本数がいろいろ変わる期待値（E_1）とに分けて考える。

また，E_1 を考える場合，いずれの層の標本数 $n_i\fallingdotseq 0$ にならない確率を P_1，その条件付きの期待値を $E_1^{(1)}$ とし，いずれかの層の標本数が０になる確率を P_2（$=1-P_1$），その条件付きの期待値を $E_1^{(2)}$ とすると，

$$E(\bar{x}')=E_1E_2(\bar{x}')$$

$$=E_1E_2\left(\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}\frac{N_i}{n_i}\sum_{j=1}^{ni}x_{ij}\Big/\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\right)=E_1\left\{\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_iE_2\left(\frac{1}{n_i}\sum_{j=1}^{ni}x_{ij}\right)\Big/\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\right\}$$

$$=E_1\left(\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\mu_i\Big/\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\right)=P_1E_1^{(1)}\left(\frac{1}{N}\sum_{i=1}^{L}N_i\mu_i\right)+P_2E_1^{(2)}\left(\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\mu_i\Big/\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\right)$$

$$=P_1\mu+P_2\mu'$$

ここで　$\mu'=E_1^{(2)}\left(\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\mu_i\Big/\sum_{\substack{i=1\\(ni\fallingdotseq 0)}}^{L}N_i\right)$

（注２） $V(\bar{x}')=E\{\bar{x}'-E(\bar{x}')\}^2=E_1E_2\{\bar{x}'-E(\bar{x}')\}^2$

$$=P_1E_1^{(1)}E_2\{\bar{x}'-E(\bar{x}')\}^2+P_2E_1^{(2)}E_2\{\bar{x}'-E(\bar{x}')\}^2$$

$n_i\fallingdotseq 0$ の場合，すなわち $E_1^{(1)}E_2\{\bar{x}'-E(\bar{x}')\}^2$ についてみると

$$\bar{x}'-E(\bar{x}')=\bar{x}'-(P_1\mu+P_2\mu')=(\bar{x}'-\mu)+P_2(\mu-\mu')$$

$$\bar{x}'-\mu=\frac{1}{N}\sum_{i=1}^{L}\frac{N_i}{n_i}\sum_{j=1}^{n_i}x_{ij}-\mu$$

$$=\frac{1}{N}\sum_{i=1}^{L}N_i\bar{x}_i-\frac{1}{N}\sum_{i=1}^{L}N_i\mu_i=\frac{1}{N}\sum_{i=1}^{L}N_i(\bar{x}_i-\mu_i)$$

$$\{\bar{x}'-E(\bar{x}')\}^2=(\bar{x}'-\mu)^2+2P_2(\bar{x}'-\mu)(\mu-\mu')+P_2^2(\mu-\mu')^2$$

$$(\bar{x}'-\mu)^2=\frac{1}{N^2}\sum_{i=1}^{L}N_i^2(\bar{x}_i-\mu_i)^2+\frac{1}{N^2}\sum_{i\neq i'}^{L}N_i(\bar{x}_i-\mu_i)N_{i'}(\bar{x}_{i'}-\mu_{i'})$$

$$E_1^{(1)}E_2\{2P_2(\bar{x}'-\mu)(\mu-\mu')\}=2P_2(\mu-\mu')E_1^{(1)}E_2(\bar{x}'-\mu)=0$$

$$E_1^{(1)}E_2\left\{\frac{1}{N^2}\sum_{i\neq i'}^{L}N_i(\bar{x}_i-\mu_i)N_{i'}(\bar{x}_{i'}-\mu_{i'})\right\}$$

$$=E_1^{(1)}\left\{\frac{1}{N^2}\sum_{i\neq i'}^{L}N_iE_2(\bar{x}_i-\mu_i)N_{i'}E_2(\bar{x}_{i'}-\mu_{i'})\right\}=0$$

$$E_1^{(1)}E_2\left\{\frac{1}{N^2}\sum_{i=1}^{L}N_i^2(\bar{x}_i-\mu_i)^2\right\}$$

$$=E_1^{(1)}\left(\frac{1}{N^2}\sum_{i=1}^{L}N_i^2\cdot\frac{N_i-n_i}{N_i-1}\cdot\frac{\sigma_i^2}{n_i}\right)$$

$$=\frac{1}{N^2}E_1^{(1)}\left\{\sum_{i=1}^{L}\left(\frac{N_i^3}{N_i-1}\cdot\frac{\sigma_i^2}{n_i}-\frac{N_i^2}{N_i-1}\cdot\sigma_i^2\right)\right\}$$

$$E_1^{(1)}\left(\frac{1}{n_i}\right)=\frac{1}{nw_i}+0\left(\frac{1}{n^2}\right)$$

$$\fallingdotseq\frac{1}{N^2}\sum_{i=1}^{L}\left(\frac{N_i^3}{N_i-1}\cdot\frac{\sigma_i^2}{nw_i}-\frac{N_i^2}{N_i-1}\cdot\sigma_i^2\right)$$

$$\fallingdotseq\frac{1}{N^2}\sum_{i=1}^{L}\left(N_i^2\frac{\sigma_i^2}{nw_i}-N_i\sigma_i^2\right)$$

$$=\frac{1}{N^2}\sum_{i=1}^{L}\left(\frac{N^2w_i^2}{nw_i}-Nw_i\right)\sigma_i^2$$

$$=\frac{N-n}{N}\cdot\frac{1}{n}\sum_{i=1}^{L}w_i\sigma_i^2=\frac{N-n}{N}\cdot\frac{\sigma_w^2}{n}$$

$$E_1^{(1)}E_2\{\bar{x}'-E(\bar{x}')\}^2\fallingdotseq\frac{N-n}{N}\cdot\frac{\sigma_w^2}{n}+P_2^2(\mu-\mu')^2$$

$$V(\bar{x}')\fallingdotseq P_1\left\{\frac{N-n}{N}\cdot\frac{\sigma_w^2}{n}+P_2^2(\mu-\mu')^2\right\}+P_2E_1^{(2)}E_2\{\bar{x}'-E(\bar{x}')\}^2$$

$$=P_1\cdot\frac{N-n}{N}\cdot\frac{\sigma_w^2}{n}+P_2[P_1P_2(\mu-\mu')^2+E_1^{(2)}E_2\{\bar{x}'-E(\bar{x}')\}^2]$$

$P_1\fallingdotseq 1$ すなわち $P_2\fallingdotseq 0$ であれば

$$V(\bar{x}')\fallingdotseq\frac{N-n}{N}\cdot\frac{\sigma_w^2}{n}$$

（注3） 単純集計の推定式 $\left(\bar{x}=\frac{1}{n}\sum_{i=1}^{L}\sum_{j=1}^{n_i}x_{ij}\right)$ の分散は次のようになる。

$$V(\bar{x})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}\fallingdotseq\frac{N-n}{N}\cdot\frac{\sigma^2}{n}$$

したがって，層別集計の単純集計に対する相対精度は，次のようにして算出される。

$$\eta=\frac{V(\bar{x})}{V(\bar{x}')}\fallingdotseq\left(\frac{N-n}{N}\cdot\frac{\sigma^2}{n}\right)\Big/\left(\frac{N-n}{N}\cdot\frac{\sigma_w^2}{n}\right)$$

$$=\frac{\sigma^2}{\sigma_w^2}=\frac{\sigma_b^2+\sigma_w^2}{\sigma_w^2}=1+\frac{\sigma_b^2}{\sigma_w^2}$$

（注4） $L=2$ の場合

$$P_r(n_i=0)=P_r(n_1=0,\ n_2=n)+P_r(n_1=n,\ n_2=0)$$

$$=(1-w_1)^n+(1-w_2)^n$$

$$=\sum_{k(1)=1}^{L}(1-w_{k(1)})^n$$

$L=3$ の場合

$$P_r(n_i=0)=P_r(n_1=0)+P_r(n_2=0)+P_r(n_3=0)-P_r(n_1=0,\ n_2=0,\ n_3=n)$$

$$-P_r(n_1=0,\ n_2=n,\ n_3=0)-P_r(n_1=n,\ n_2=0,\ n_3=0)$$

$$=(1-w_1)^n+(1-w_2)^n+(1-w_3)^n-(1-w_1-w_2)^n-(1-w_1-w_3)^n$$

$$-(1-w_2-w_3)^n$$

$$=\sum_{k(1)=1}^{L}(1-w_{k(1)})^n-\sum_{k(1)=1}^{L}\sum_{k(2)\neq k(1)}^{L}(1-w_{k(1)}-w_{k(2)})^n$$

上記のようにして，一般の $P_r(n_i=0)$ の算式が成り立つことが類推できるであろう。

（調査部　経済統計課）

Stratified Tabulation of the Unincorporated Enterprise Survey

Mitsuru Ide

The results of the Unincorporated Enterprise Survey contain a large sampling error, because the sample size is not large enough. There are, however, many reasons why the enlargement of the sample has not been implemented.

In the paper, the author discusses one of the methods which makes the sampling error small without the enlargement of sample size. This method may be called "stratified tabulation" tentatively, because the variance of estimation formula is approximate to that of stratified sampling (proportional allotment).

The tabulation method is as follows:

Firstly, the average of each stratum is calculated, when the sample enterprises are stratified according to their industry and the number of employees. In this operation the distribution of enterprises by industry and number of employees is obtained from the Establishment Census. Finally, the total average is calculated as a weighted arithmetic mean of the above-mentioned averages by using the distribution from the Census.

The method includes two biases. Namely, one is occurred from non-existence of sample enterprise found in some strata and another from the use of distribution obtained from the other source which may be three years old.

However, when the strata are made resonably, the results experimentally calculated from the data of the previous years show that the biases are considerably smaller than the sampling error. Accordingly, the author proposes the application of the method.

住民基本台帳にもとづく人口移動報告の一問題点について

井 上 俊 一
高 橋 邦 明

目　　次

は じ め に

住民基本台帳法（昭和42年，法律第81号）によれば，人がその住居を変更するには，まず旧住所地の市町村役場におもむき転出の届出を行ない，転出証明書の交付をうけ，転出の後は，新住所地の市町村役場においてその転出証明書を添えて転入の届出をするものと定められている。この制度にもとづいて人口移動統計を作成するには，したがって，理論的には転出届によって統計を作成する方法，転入届によって統計を作成する方法，およびその両者を併用する方法が考えられる。しかし，昭和42年11月以前には，転出届，転出証明書の制度