

RESEARCH MEMOIR OF
THE STATISTICS BUREAU

March 1994

No. 52

CONTENTS

《Articles》

Some Issues of the Calculation of Purchasing Power Parities
and Real GDP in the International Comparison Programme
Tatsuya Ono ... 1(41)

Examination of improvement in sample design on the
Unincorporated Enterprise Survey Kunihiro Koike ... 43(67)

《Introduction》

Design Technique of Statistical Information Database
—for performance improvement— Fumihiko Nishi ... 69(93)

《Note》

On Suppression of Statistical Data Katsutoshi Nagashima ... 95(116)

STATISTICS BUREAU AND STATISTICS CENTER
MANAGEMENT AND COORDINATION AGENCY
JAPAN

参考4-2

ISSN 0446-5849

統計局研究彙報

1994年3月
第52号

統計局研究彙報

52

《論文》

国際比較プログラムにおける購買力平価及び 小野達也 ... 1
実質GDP算出上の諸問題について

個人企業経済調査標本設計における改善の検討 小池邦彦 ... 43

《紹介》

統計情報データベースの設計技法 西文彦 ... 69
—パフォーマンス向上のための技法—

《記録》

統計データの秘匿について 永島勝利 ... 95

総務庁統計局 統計センター

総務庁統計局・統計センター

個人企業経済調査標本設計における 改善の検討

小池 邦彦

目 次

はじめに

I 標本交替方式

II 最近の結果の標本誤差

III 現行方法による問題点

IV 改善方策

V 改善方策の検証

おわりに

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、昭和27年に統計局に移管された。この調査の初期の目的は、国民所得推計のための基礎資料を作成することであったため、標本数も1,000事業所と極めて少なかった。移管後は、個人企業の経営の実態及び時系列変化を明らかにすることとなってきたため、漸次、標本数を増やし、現在は約2,800企業となっている。しかし、四半期別結果の時系列的不安定性が残り、まだ十分な標本数とはいえ、このため、比推定の考え方をとり入れた従業者規模修正を行い、売上・収益等の実額は単純推計により算出し、対前年度（同期）の増減率は従業者規模修正を用い結果精度の向上を図っている。

しかし、この従業者規模修正値は、単純推計値から求めた増減率と不整合があり、利用者が理解しにくい場合がある。（図1）

また、従業者規模修正により結果の安定性は増すにしても、単純推計の結果精度が必然的に従業者規模修正の結果に影響することはいうまでもない。こうした点を勘案すると、単純推計の結果の安定性を図ることが不可欠となる。

以下、最近の結果等を見ながら、不安定性の要因と思われるものを探り、実際に改善できる点について検討を行った。

I 標本交替方式

昭和55年度以降の産業別売上高（図2）をみると、62年度以前と63年度以降で明らかな違いがある。この変化の原因と考えられるものとして、この間の標本交替方式の切替えが考えられる。

この切替えは、単位区の交替による変化を極力平準化するために行われたものである。

1 昭和62年度以前（2交替方式）

単位区をA、B 2つの組に分けて1期（四半期）ずつ交替する方式をとっていた。単位区は1年継続して調査を行うが、調査事業所は前半の2期と後半の2期で交替する。

この際、大規模事業所（従業者数10人以上の製造業）については、単位区内の事業所数が極めて少ないこと及び結果数値に与える影響が大きいことから1年間継続して調査することと、できるだけ前半に大規模事業所が抽出されるよう起番号を設定した。このため、前半の抽出事業所の従業者規模が後半の抽出事業所の従業者規模より大きくなっていった。

標本単位区及び調査事業所の交替状況（62年度以前）

交替区分		1期	2期	3期	4期	1期	2期	3期	4期
A	単位区	○				○			
	事業所	●	●	▲	▲	●	●	▲	▲
B	単位区		○				○		
	事業所	▲	●	●	▲	▲	●	●	▲

●：前半6か月分の抽出事業所の従業者規模 > ▲：後半6か月分の抽出事業所の従業者規模

この結果、図のとおり期別の事業所の従業者規模が、第2期は●が集中し、第4期は▲が集中していたため、売上高等の結果数字にも影響を与えていた。

2 昭和63年度以降（4交替方式）

全体の単位区をA、B、C、D 4つの組に分け、每期1組を交替する。単位区は1年継続して調査を行い、調査事業所は前半の2期と後半の2期で交替する。前半と後半の事業所の選定は、昭和62年度以前と同様であり、前半の抽出事業所の従業者規模が後半より大きくなっている。

期別の事業所の従業者規模をみると、●の前半事業所、▲の後半事業所が各期とも2つずつとなり、期別の抽出事業所の従業者規模は同一となっている。

標本単位区及び調査事業所の交替状況 (63年度以降)

交替区分		1期	2期	3期	4期	1期	2期	3期	4期
A	単位区	○				○			
	事業所	●	●	▲	▲	●	●	▲	▲
B	単位区		○				○		
	事業所	▲	●	●	▲	▲	●	●	▲
C	単位区			○				○	
	事業所	▲	▲	●	●	▲	▲	●	●
D	単位区				○				○
	事業所	●	▲	▲	●	●	▲	▲	●

●：前半6か月分の抽出事業所の従業員規模 > ▲：後半6か月分の抽出事業所の従業員規模

II 最近の結果の標本誤差

個人企業経済調査報告書には、売上高の標準誤差率は、約6%と記載されているが、62年度以前の2交替方式及び63年度以降の4交替方式による売上高について単純任意抽出法によった場合と仮定して次式により標本誤差を算出した。

1 単純推計値の標本誤差

$$\sigma(\bar{x}) = \frac{\sqrt{(N-n)} \cdot S}{\sqrt{N} \cdot \sqrt{n}} \quad \left\{ S^2 = \frac{\sum f(X - \bar{X})^2}{\sum f} \right\}$$

f：線型推定乗率

n：線型推定乗率付与済企業数

2 規模修正値の標本誤差

$$M.S.E(R) = R^2 \cdot \frac{N-n}{N-1} \cdot \frac{C_x^2 - 2\rho C_x C_y + C_y^2}{n}$$

R：従業者1人当たりの売上高 (=X/Y)

X：売上高 (単純推計値)

Y：従業者数 (単純推計値)

C_x：変動係数 = σ_x / \bar{X}

ρ：XとYとの相関係数

C_y：変動係数 = σ_y / \bar{Y}

表1 期別売上高の標準誤差率 (昭和62年度1期～4期の平均)

	製造業	卸・小、飲	サービス業
単純推計	5.68	4.06	5.55

表2 期別売上高の標準誤差率及び従業者数との相関係数 (平成元年度1期～4年度4期の平均)

	製造業	卸・小、飲	サービス業
単純推計	5.47	3.90	4.93
規模修正	4.19	4.43	3.30
相関係数	0.65	0.43	0.78

この結果、表1と表2の単純推計値を較べると、2交替方式から4交替方式に変更したことにより、結果数字の変動幅は、3産業とも小さくなったことがわかる。

表2の結果をみると、製造業及びサービス業については、規模修正の効果があるが、卸売・小売業、飲食店については、単純推計値より変動幅が大きい。

これは、卸売・小売業、飲食店の売上高と従業者数の相関が小さいためである。

また、卸売・小売業、飲食店の売上高を時系列でみても、単純推計値と規模修正の結果にさほど差がない。(図1)

これらの点から、卸売・小売業、飲食店に限れば、規模修正を行わず、単純集計結果のみで公表することも可能であると思われる。

Ⅲ 現行方法による問題点

1 組別の従業者規模の不均一の問題

標本交替方式を2交替方式から4交替方式に変更したことにより、期別の結果数字の変動幅は縮小したが、調査が小規模等の要因により、4組全体の事業所の従業者規模は、依然として不安定性が残っているのが現状である。

つまり、「1 標本交替方式」の「2 昭和63年以降」で示したとおり、前半抽出事業所の従業者規模>後半抽出事業所の従業者規模となっており、期ごとの従業者規模は、前半×2組、後半×2組で、差異はないはずである。

しかし、実際は、4組ごとの調査市区町村が異なり、同じ大規模事業所でも従業者規模は異なる。例えば、第1期をみると、A組とD組は、前半6か月の調査対象で従業者規模が大の●であるが、事業所の従業者規模は調査市町村が異なるため、同じとはいえない。

2 母集団の陳腐化の問題

個人企業経済調査の母集団フレームは事業所統計調査結果であり、事業所統計調査の情報から個人経営の事業所を抽出して調査を実施している。このため、フレームが5年間据え置きとなり、この間の事業所の変動(事業所の業種変更、移転、法人化等)が反映されない。

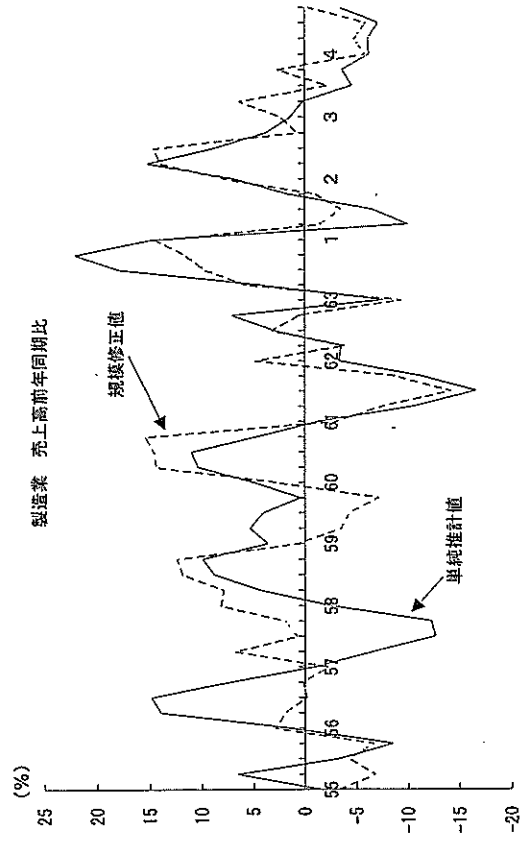
この結果、当初調査予定事業所に調査対象外となる不適格が多くなり、代替として抽出される事業所の従業者規模は、小さくなる傾向にある。(調査事業所

である。
値と規模
が、単純

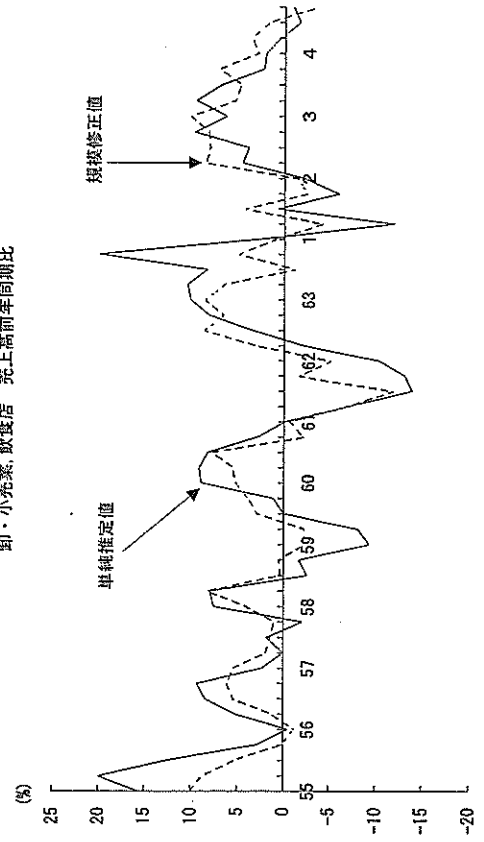
別の結
の事業
前半
月ごと
でも
月の
村が

新統
め、更
替所

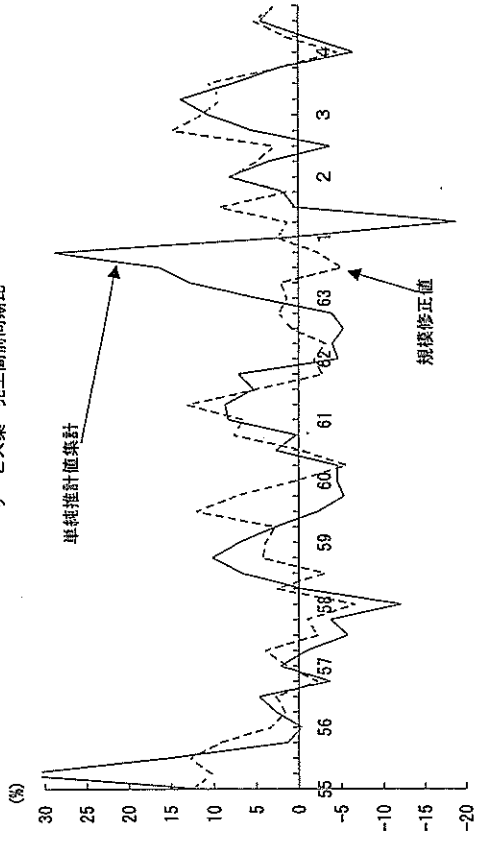
図1 産業別売上高対前年同期比



卸・小売業、飲食店 売上高対前年同期比



サービス業 売上高対前年同期比



の抽出は、調査単位区内の調査対象事業所を従業者数の大きい順に配列し、所定の抽出起番号及び抽出間隔で系統抽出する方法によって行っているが、代替事業所を抽出する際は、その事業所の前後の事業所としている。）

なお、この傾向は、実際の調査拒否・調査不能によっても生じる。

3 単位区の設定上の問題

個人企業経済調査の単位区は、調査員の活動範囲と単位区内の総事業所数(所定事業所数の確保)を考慮して設定するため、事業所密集地域つまり繁華街地域が単位区となるケースが多い。このため、繁華街から離れた場所に立地していると考えられる製造業の事業所(特に大規模事業所)が単位区内に存在する割合が少ない。

この結果、製造業の事業所数が少なく推計される。

4 層別抽出の問題

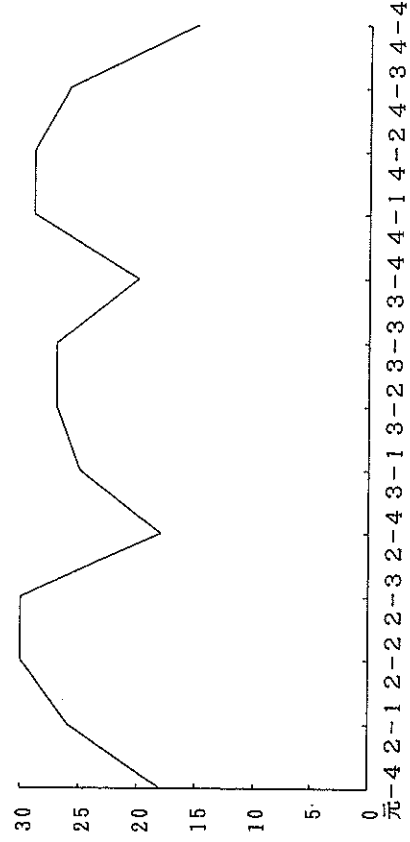
個人企業経済調査は、層別抽出を行っているが、その層化基準は、市町村の製造業事業所比率であり、地方区分等により、全体を194層に分けて1層から1市町村1単位区を抽出している。

しかし、予備企業の確保、都道府県別調査市町村数の固定という点から、単位区を設定できない層が毎年度存在し、期別の構成層をみると、重複又は欠如している層が4期を除き、多く発生している。これは、年度別の設定層が異なっており、4期以外は、前年度設定の交替区分と当年度設定の交替区分が混在しているためである。

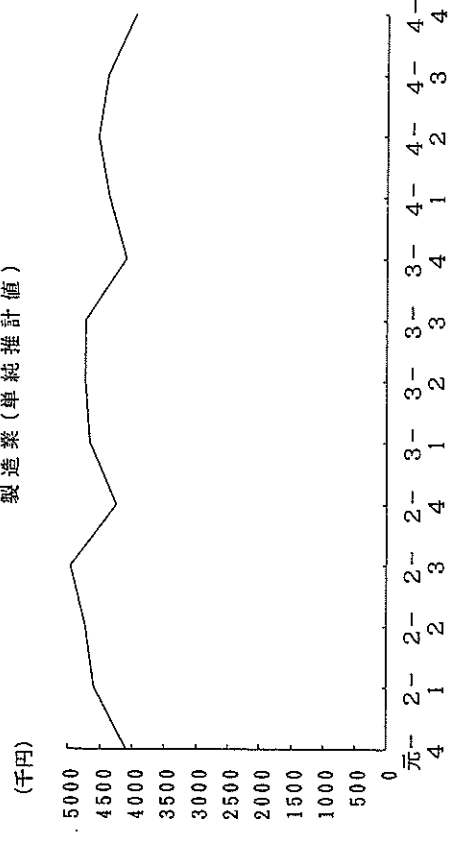
この重複又は欠如層数の差が、図2に示すとおり、期別の売上高に影響を与えていると思われる。

図2 期別重複・欠如層数及び産業別売上高

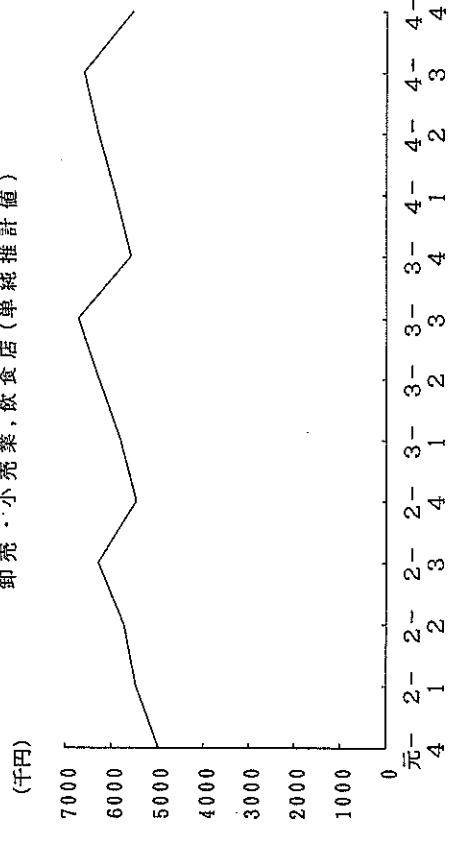
重複・欠如層数



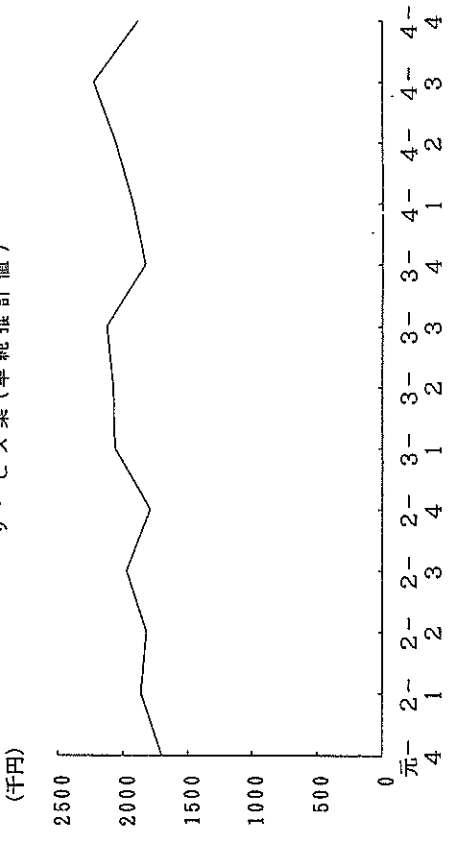
製造業(単純推計値)



卸売・小売業, 飲食店(単純推計値)



サービス業(単純推計値)



Ⅳ 改善方策

1 起番号のランダム化

単位区内の抽出事業所を前半抽出事業所の従業員規模>後半の抽出事業所の従業員規模となるように選定する方法は、起番号を前半抽出の起番号<後半抽出の起番号となるよう起番号の操作を行っている。

そこで、この起番号の操作を行わず、前半と後半の抽出事業所の従業員規模を均一にした方が期別の従業員規模の差は少なくなると思われる。

起番号をランダム化することにより、下図のとおり、組別の従業員規模は異なるものの期ごとを通じ、◆■★▼の4種類が揃い、全体の従業員規模は一定となる。

交替区分	1期	2期	3期	4期	1期	2期	3期	4期
A 単位区 事業所	◇	◇	◇	◇	◇	◇	◇	◇
B 単位区 事業所	■	■	■	■	■	■	■	■
C 単位区 事業所	★	★	★	★	★	★	★	★
D 単位区 事業所	▼	▼	▼	▼	▼	▼	▼	▼

2 母集団の切替え

現行では、事業所名簿整備（新設、廃業）の結果を抽出名簿作成時に、母集団リストとして反映しているが、母集団フレームの陳腐化を防ぐため、母集団切替えについても、事業所名簿整備を悉皆調査で行い、さらに経営組織も調査

して利用可能な調査とし、切替え周期を短縮する。

3 単位区設定方法の変更

1事業所当たりの従業者数を平成3年事業所統計調査の結果と比較すると、次のように個人企業経済調査の結果は、各産業とも下回っており、特に製造業は、その差が大きい。これは、売上高と従業者数の相関が高い製造業にとっては、売上高が過少に推計されることになる。

1事業所当たりの平均従業者数 (人)

	製造業	卸・小売業	サービス業
平成3年事業所統計調査	3.65	2.91	2.30
個人企業経済調査結果	3.08	2.74	2.18

この原因は、調査単位区の抽出の際、製造業事業所の従業者規模の小さい単位区が調査対象となっているためである。

調査単位区を設定する際、3産業の事業所数がある程度の大きさになるよう設定するが、大規模製造業が立地すると考えられる郊外は、3産業の事業所数が少ないため、選定されにくい。

つまり、卸売・小売業、飲食店及びサービス業が集中する市街地の単位区には、大規模製造業が少ないため、製造業の平均従業者数が他の産業に比べ大きく落ち込むことになっている。

この問題を解決するため、卸売・小売業、飲食店及びサービス業については、現行の単位区設定方法により市街地に設定し、製造業については、新たに別の単位区を市街地から離れた地域に設定する。その際の3産業別の調査事業所数は、当該市町村の修正従業者数の割合により配分する。

4 層化基準の見直し

層の重複・欠如は、個人企業の多い市町村の代わりに個人企業の少ない市町

村が抽出されていることを示している。市町村内の個人企業数が、1企業当たりの売上高にどのように影響するかは、地域の特性によって異なると考えられるが、期別の重複・欠如層数を少なくするためには、年度毎の構成層及び組別の構成層を同一にする必要がある。

このためには、現行の層を粗くすることが考えられる。

V 改善方策の検証

1 単位区内抽出起番号のランダム化

起番号の置き換え操作を行わず、全くのランダムで起番号を発生させた場合の1事業所当たりの従業者数、売上高を実際のデータを組み合わせて試算した。

起番号をランダム化した場合の数値を算出するため、期別、組別のデータを企業一連番号の奇数組・偶数組の2つのグループに分け、それぞれをランダム化した場合の組み合わせに置き換え、そのデータを期別に集計する。

起番号ランダム化による売上高の期別標準誤差率は、元のデータが同じであるので、期別結果を平均すると現行の方法によるものと同じになるが、時系列のグラフをみると、各産業とも前期との差が小さくなり、なだらかな曲線となる。(図3)

しかし、ランダム化の組み合わせは、調査時期の異なるデータを合わせて集計するものであり、売上高等の季節的変動の大きい結果は、時期的なズレが影響し、正確な比較ができない。そこで、季節的変動の少ない1事業所当たりの従業者数について比較する。

その結果を時系列のグラフでみると、各産業特に製造業及びサービス業は、ランダム化した結果の方がなだらかな曲線となっている。(図4)

2 事業所名簿整備のデータ利用

次回(平成6年)の事業所名簿整備では、「経営組織」が調査事項となってい

図3 産業別売上高

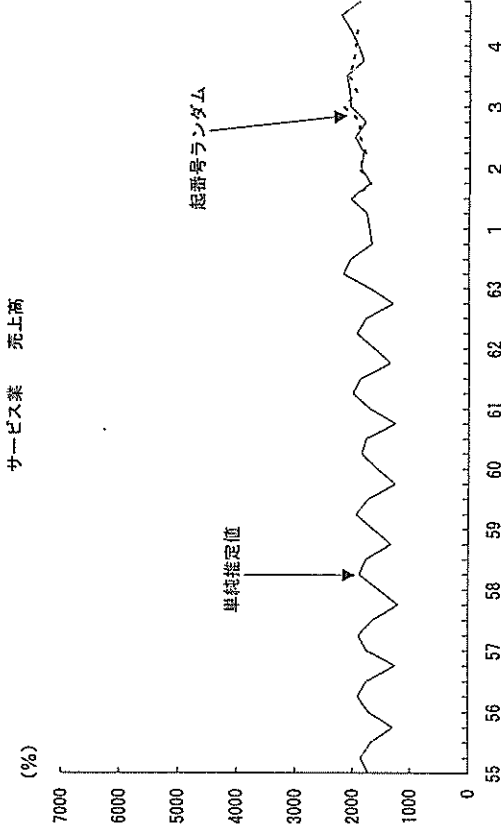
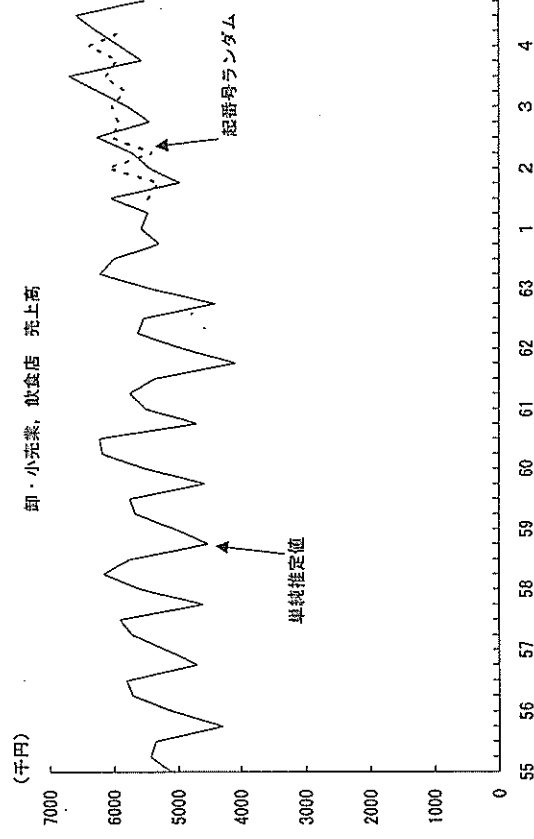
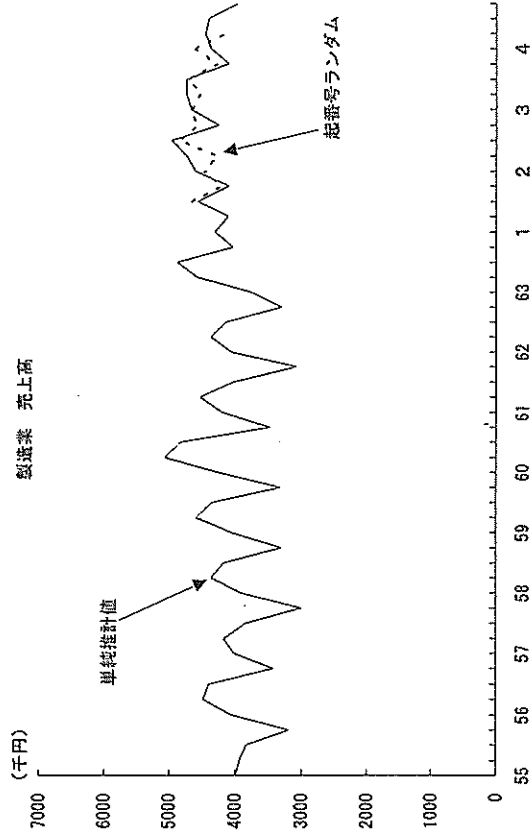
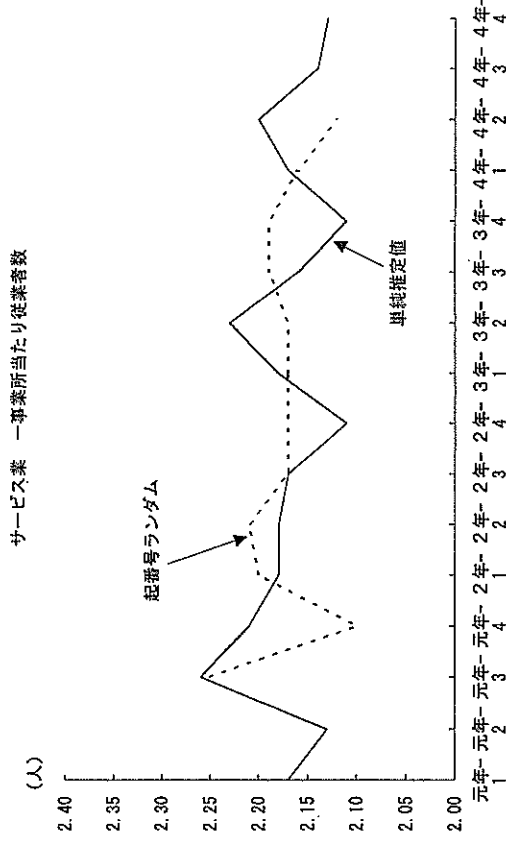
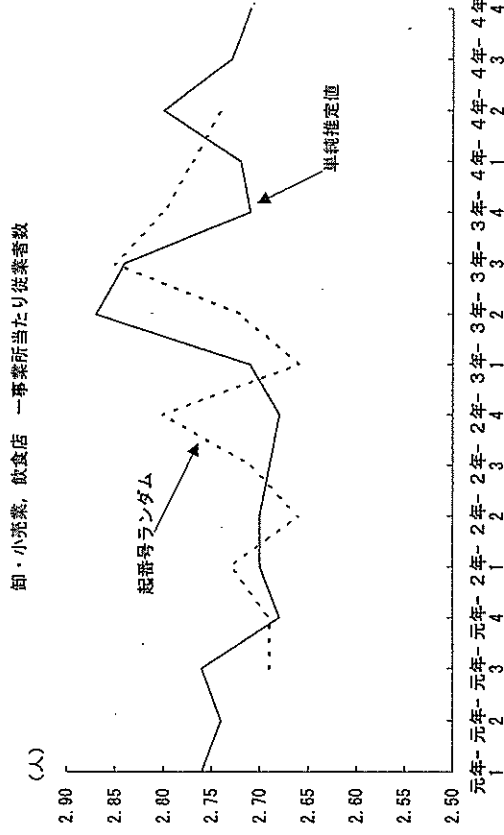
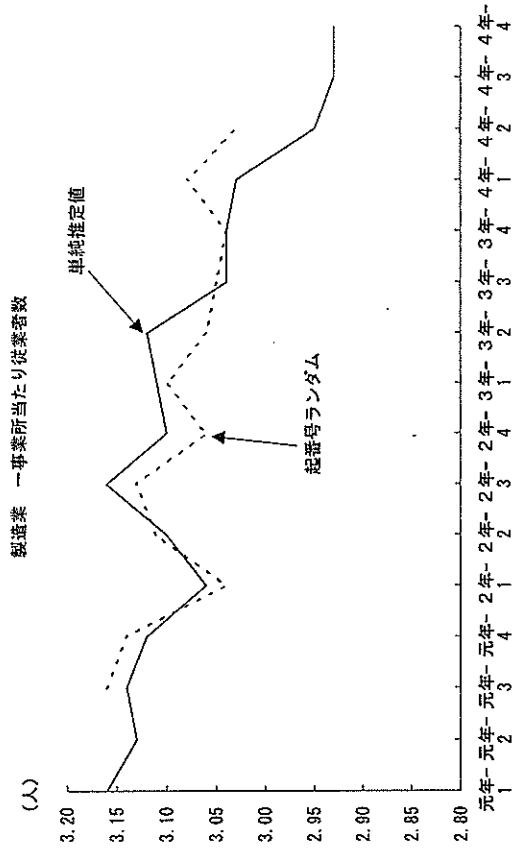


図4 産業別一事業所当たりの従業員数



ないため、個人経営の事業所を限定できないが、次回以降の事業所名簿整備において、調査事項として「経営組織」を加え、事業所名簿整備の結果を母集団リストとして利用する。

3 1市町村2単位区設定方式

現行の方式は、1市町村1単位区を設定し、単位区内の3産業別調査事業所数の配分は、単位区内の修正従業者数の割合により決定しているが、市街地に設定した単位区は、市町村全体の縮図になっているとは限らない。

そこで、市町村全体の修正従業者数の割合から、3産業別の調査事業所数を配分し、平成6年度の調査対象市区町村について、製造業の配分値が確保できるよう試験的に製造業単位区を設定し、製造業単位区の面積を測定した。

調査単位区的面積

(km²)

	市 部	町村部	全国平均
現行の方式	2 以内	4 以内	2.4
2 単位区方式	8.3	22.7	15.1

その結果、個別の市区町村をみると、現行方式の面積以内に収まる市区町村もあるが、平均では1市2単位区方式の面積は、現行方式の4～6倍となり、調査員の稼働面積を考えると実行困難と思われる。

おわりに

層化基準の見直しについては、層化した単位区の売上高及び1事業所当たりの従業者数を3産業別に算出し、層化基準である製造業事業所比率との関係を探ったが、明らかな関係は見当たらなかった。このため、製造業事業所をなるべく多く抽出するという現行の層化基準の目的を崩してまでも、層数を減らすという新たな層化を行うべきであるという結論には到らなかった。

また、期別の結果を安定させるための起番号のランダム化は、実現可能であ

るが、製造業の調査事業所数の過少を解決するための1市町村2単位区方式は、全市区町村についての確認が必要となり、また、調査経費、調査員の選任等の問題が残る。

個人企業経済調査標本設計における改善の検討

個人企業経済調査の初期の目的は、国民所得の推計のための基礎資料を得るためのもので、標本誤差もあまり問題にならなかったが、その後、個人企業の経営の実態及び時系列変化を明らかにすることが目的となり、何度か標本数の拡大を行い、現在は194単位区、約2800企業を調査している。しかし、まだ十分な標本数とはいえず、このため、比推定の考え方をとり入れた従業者規模修正を行い、売上・収益等の実額は単純推計により算出し、対前年度（同期）の増減率は従業者規模修正を用い結果精度の向上を図っている。

しかし、この従業者規模修正値は、単純推計値から求めた増減率と不整合があり、利用者が理解しにくい場合がある。

また、従業者規模修正により結果の安定性は増すにしても、単純推計の結果精度が必然的に従業者修正の結果に影響することはいうまでもない。こうした点を勘案すると、単純推計の結果の安定性を図ることが不可欠となる。

そこで、最近の結果等を見ながら、不安定性の要因と思われるものを探り、改善できる点について検討を行った。

- (1) 標本交替方式を2交替方式から4交替方式に変更したことにより、期別の結果数字の変動幅は縮小したが、組別の事業所の従業者規模は調査市町村が異なるため同じとはいえない。

そこで、調査事業所の抽出には、起番号をランダム化することが考えられる。

- (2) 個人企業経済調査の母集団フレームは事業所統計調査結果であり、フレームが5年間固定になる。

そこで、母集団フレームの陳腐化を防ぐため、事業所名簿整備の結果を利用することが考えられる。

- (3) 個人企業経済調査の調査単位区は、調査員の活動範囲と単位区内の総事

業所数を考慮して設定するため、繁華街地域が単位区となるケースが多く、製造業の事業所が抽出される割合が少ない。

そこで、単位区の設定方法を1市町村2単位区に変更することが考えられるが、実現するには問題が残る。

- (4) 個人企業経済調査の層化基準は、市町村の製造業事業所比率であり、都道府県別調査市町村数の固定という点から毎年度単位区を設定できない層がある。

そこで、每期同じ層から抽出されるように、層を粗くし、毎年度の構成層を同一にすることが考えられる。

Examination of improvement in sample design
on the Unincorporated Enterprise Survey

Kunihiko Koike

The initial purpose of the Unincorporated Enterprise Survey was to get the basic material data for estimate of National income. But now, the survey aims at collecting data on the activities and financial status of unincorporated enterprises and the changes in time series. Though about 2,800 enterprises in 194 sample unit districts have been covered in this survey, the sample size is not large enough. Therefore, in order to improve the accuracy of results, we have adjusted the number of persons engaged with the ratio estimate, calculated the real value of revenue or sales volume with the simple estimate, and adjusted the change over the year by using the adjusted number of persons engaged. It is, however, difficult for users to understand why the increase rate over the previous fiscal year or term calculated by the adjusted number of persons engaged is not consistent with that calculated by the simple estimate. The reliability of the results is improved by the adjustment of the number of persons engaged, whereas the adjustment of that is affected by the degree of accuracy of results by simple estimate. Therefore, it is indispensable to improve the reliability of results by simple estimate. We examined the causes of unreliability. The results are as follows.

- 1) Quarter of sample enterprises change each survey period, so accuracy of the results each of period is improved. However, the size of sample enterprises is different each period because the sample size is not large enough.

Then, the random sampling should be used in the selection of enterprises for the survey.

- 2) The frame of universe of the Unincorporated Enterprise Survey is

set up by the results of the Establishment Census. So, it is fixed for 5 years.

The establishment directory should be used for setting up the frame in order to prevent it from being out of date.

- 3) Many of sample unit districts are business quarters and there are a few manufacturing enterprises in these districts. This is because the sample unit districts are selected based on the number of enterprises in the district which an enumerator can cover.

Then, two sample unit districts should be selected in a municipality. But some problems should still remain for the materialization of it.

- 4) Sample municipalities are divided into groups based on the proportion of manufacturing enterprises in a municipality. But, in some groups, sample unit districts can not be selected every year.

Then, it should be considered decreasing the number of groups so as to select the sample from the same group each period.

統計情報データベースの設計技法

—パフォーマンス向上のための技法—

西 文 彦

目 次

はじめに

- I 旧 SISMAC におけるパフォーマンス上の問題点
- II 新 SISMAC におけるパフォーマンス改善のための技法
- III パフォーマンスにおける改善度の評価

おわりに