日本に於ける法人企業の標本抽出法について*

――ある不等確率抽出法における最適抽出確率に関する一考察――

田　口　時　夫

(1974年10月受付)

On the Optimal Probability of Selection in Some Unequal Probability Samplng

—On a Sampling Design For the Corporation Survey in Japan—

Tokio Taguchi

In this paper, the author examines what kind of sampling system for the corporation survey is the most suitable from the theoretical and empirical point of view.

In Section 3, procedures of two simple random samplings with replacement, one with equal probability of selection, where the estimate using sample mean and the ratio estimate are considered, and another with size proportional probability of selection are studied under the assumption that the population distribution is the Mardia-Pareto bivariate distribution of the 1st kind with respect to X and Y.

In this case, the samplng errors of both methods can be compared through $C=\frac{\sigma_x}{\mu_x}$, $C'=\frac{\sigma_x'}{\mu_x}$ and $R^2=\frac{\sigma_x'^2}{\sigma_x^2}$, where μ_x is the population mean, σ_x^2 the population variance and $\sigma_x'^2=\mu_y E\left(\frac{X^2}{Y}\right)-\mu_x^2$, where X and Y represent the main and supplemental variable respectively.

Particularly, for the Mardia-Pareto distribution of the 1st kind, C, C' and R^2 are all decided only by Pareto coefficient α and the correlation coefficient between X and Y is given by $1/\alpha$.

From these relationships, the author induces some criteria choosing sampling method.

In Section 4, the two stratified random samplings with replacement, with equal and size proportional probabilities of selection and with strata constructed as usual with respect to some conditions of the supplemental variable Y, are considered from the same points of view as in Section 3. It is observed that the conventional stratification of enterprises is more efficient for the samplng with size proportional probability of selection than for the sampling with equal probability of selection.

In Section 5, the optimality of probability of selection on the Pareto population is discussed. Provided to the probability of selection proportional to Y^ν; $\nu<\alpha$, $\alpha+\nu>2$, the optimal probability of selection $\nu^{(0)}$ is only decided by α and satisfies $|\nu^{(0)}|<\alpha$ (simulatively $0<\nu^{(0)}<1$) for the above Pareto distribution.

The minimum error of this kind of sampling with varing probability of selection is studied at the same time.

In Section 6, the Y^ν proportional probability sampling is considered on Gibrat population. And it results that if $Y^{\nu^{(0)}}$ is the optimal probability of selection, $\nu^{(0)}$ is equal to logarithmic regression coefficient of X on Y and the variance of estimator is decided by sample size and the variance of its logarithmic residnals in this case.

In Section 7, the criteria obtained in the previones sections are examined based on some empirical data. It is observed that the conventional correlation coefficient is not enough to give a common criterion deciding sampling method to all cases which is treated theoretically or

* 本研究の一部は科学研究費補助金によるものである.

empirically in this paper but the logarithmic regression coefficient, logarithmic correlation coefficient and some allied measurements can be expected to be more effective in the decision of sampling method.

And from the above results, some new methods of stratification and sample allocation are suggested. In the final section, a sampling system for the corporation survey which is a mixture of the samplings with equal and size proportional probability of selection is proposed for actual use.

The Institute of Mathematical Statistics

1. はしがき（目的及び意義）

本稿は，大蔵省法人企業統計調査における標本設計の改訂の為の分析である．

従来，我国における法人企業を対象とする標本設計に於ては，等確率非復元抽出を原則とし，目標精度に達するように層化基準及び抽出率を決定するのが建前であった．

然し乍ら法人企業に関する各種の規模分布は，第1図に示されるように，正規分布に代表される対称型分布と異なり，大きな歪みを伴うものであり，その結果，大規模階層に悉皆抽出法を適用しても，尚且，特に階層別集計に関して大きな抽出誤差を生ずるのが実状である．

それ故に Hansen，M.H. 及び Hurwitz，W.N [1] は，アメリカにおける労働力調査に不等確率抽出法の適用を提案したのであるが，それは等確率抽出を鉄則とした従来の設計の立場からすれば画期的といえるものであった．

その後，任意確率抽出法は，Horvitz，D.G. 及び Thompson，D.J. [3] を初めとして理論的に整備究明を重ね今日一応の完成をみたといえる．一方，その方法は例えばインドの農業統計（The method of forcasting acreage of principal crops，c.f. Sukhatme，P.V. [7]）やギリシャの労働力調査（c.f. Raj，D [5]）にみられる如く統計実務に適用され，具体的に検討されているが，その例は尚少なく，特に我国の主要統計に於ては前例をみないのである[1)].

従って，ここに大蔵省法人企業統計に関して，その方法の検討を加えることは，統計調査上少なからぬ意義を有するものと思われる．

然し乍ら，周知のように，非復元任意確率抽出法に関しては，理論的に体系化されているとはいえ，特に大標本の場合の精度の推定に関しては煩雑な計算を必要とし（c.f. Rao，J.N.K [5]，etc）他方経済諸量の分布に関しても，その特性化は必ずしも一義的でない段階にあるから，其等の詳細については次の機会に検討するのが妥当であろう．従って，以下の所論は差当り復元等確率及び規模比例確率抽出の比較から出発し層化法，不等確率抽出法を逐次パレード及びジブラ母集団に関して思考実験或いはシミュレーションを行ない，其等の結果を経験資料に照して実証する過程を辿ったものである．結論からいえば，こうした典型の考察は，方法の適用条件及び効果について，簡潔な見通しを与えることにより理論の形成と方式の撰択に有効であると信ずる．

2. 問題の設定と基本的諸条件

一般に，理論的に可能性を検討する段階に於ては，一般的抽象的前提に立つ事が必要であるが，実践的に具体的決定を行う段階，或いは経験資料を判定する実際的立場に於ては，寧ろ典型的例題を設定し，精密に分析する方が屢々有効な指針を与えることが出来る．

---

1) 物価指数作製の際に指定される品目は取引高が一定金額を超えるものとされているが，此の方法を一般化すると規模比例抽出法に導かれる．現行に於て取引高の低い品目を無視している処に問題の一端が見出される．

(A) 母集団の分布について

所得，資用等に関する代表的な分布としては，ジブラ分布とパレート分布が最も頻煩に用いられるが，特に後者は，一定規模以上の階層に対して適合性が高いことが認められている（第1図参照）．パレート分布をジブラ分布に優先させる他の理由は，もしある二種の経済量がそれぞれジブラ分布に従うとするならば，其等の間の比率もジブラ分布に従うべきであり，其の場合例えば第1図及び第2図は同一の分布であるとせねばならない[2]が，多くの実際例によると，これは極めて無理な仮定であるといわざるを得ないからである[3]．

然し乍ら補助変量を加えた二変量に関するパレート分布については，Mardia，K. V [4] の第1種及び第2種の分布を数えるに過ぎない．此の場合，第2種分布は比較的複雑な形態をもつ反面，基本的な性格は，第1種に類似したものと思われる．従って以下の所論は，専ら次の仮設に拠ることにする．

仮説 1 「母集団分布は差当り Mardia の第1種パレート分布

$$f(x,\ y)=\frac{\alpha(\alpha+1)\,\theta_1{}^{\alpha+1}\,\theta_2{}^{\alpha+1}}{(\theta_2\,x+\theta_1\,y-\theta_1\,\theta_7)^{\alpha+2}}; \tag{1}$$

$$0<\theta_1\leqq x\leqq\infty,\ 0<\theta_2\leqq y\leqq\infty,\ \alpha>1.$$

に従うものとする[4]．但し第6節に於てジブラ母集団を対象とした.」

(B) 推定の対象について

一般に標本設計の妥当性は，推定の対象や推定の方式に依存する．

法人統計に於て，基本的に推定目的とされるものは

a) 各項目に関する総量

b) 名項目間の比率

である．a) に関しては各項目に関する平均の推定としても差支えないであろう．b) に関しては，調査対象ごとに算定した項目間比率の平均[5]によるか，項目別総量間の比率によるかによって異なるが，現行の財務比率の算定方式によれば，後者を目的としているといえる．且その場合の推定精度は，(1) の精度に大きく依存するから，差当り次のように目的を限定することが出来る．

仮設 2. 「推定対象は，母集団平均値である.」

(c) 抽出形態と推計方式について

法人企業に対しては，その大量性と多様性によって，通常大標査論が適用されている．

又，仮設2に対しては，等確率抽出法に限定しても，標本平均による他，比推定，回帰推定等の諸方式が考慮されうるが，仮説1のような母集団に対する既成の観点からみて，推計方式を次のように限定するのが先決である．

仮設 3. 「仮設2の推計方式は，復元等確率抽出（以下 ep sampling と略称する）による標本 $X_1, \cdots, X_n$ の平均による推定

$$\bar{X}=\frac{\sum_{i=1}^{n}X_i}{n} \tag{2}$$

2) 後述 (1) のパレート分布を仮定した場合比 X/Y の分布は XY 間に相関がある時単峰型歪対称分布であり，又独立のときパレート分布であることが拙稿 [7] によって証明される．

3) 更にジブラ法則の実態からの乖離は其等の集中曲線に於て生ずる．すなわちジブラ法則が成立するならば，その集中曲線は，均等線と異なる対角線に関して対称型を保つべきであるが，この仮定も多くの実例に於て成立しない．

4) 第4章に於ては，此の分布を更に補助変量 y により層別した場合を扱った．

5) 第2図によると此の比率の分布は単峰型である．

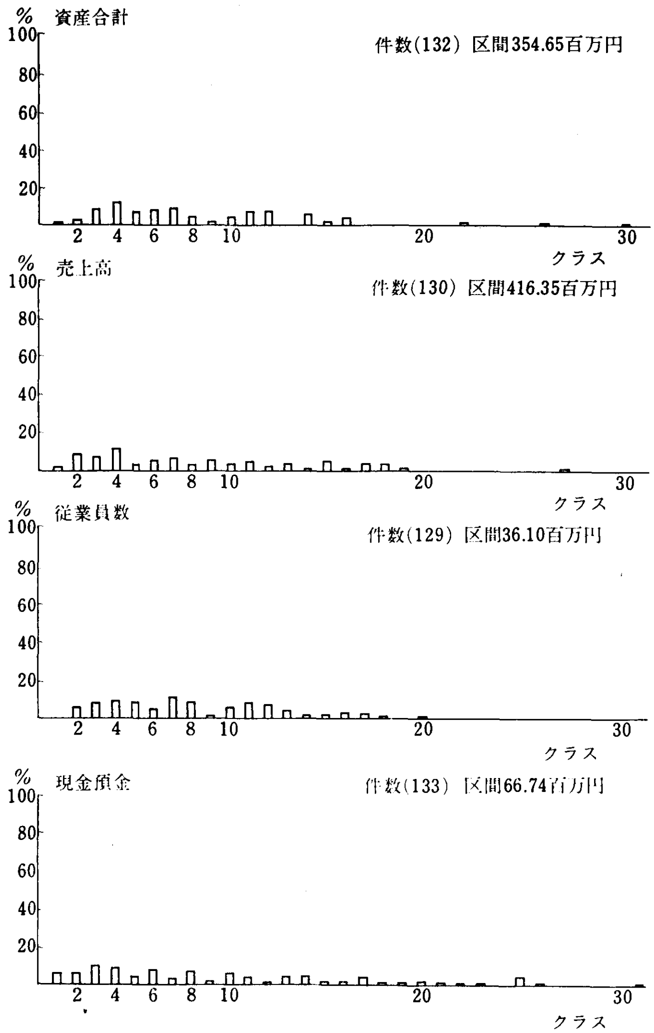
と，補助変量 y を用いた復元規模比例確率抽出（以下 ppes sampling と略称する）による標本 $(X_1', Y_1')\cdots\cdots(X_n', Y_n')$ による平均比率を用いた推定

$$\tilde{X} = \mu_y\left(\sum_{i=1}^{n}\frac{X_i'}{Y_i'}\right) \tag{3}$$

に限定する[6]，但し μ_y は y の母集団平均を表わすものとする．すなわち $\mu_y = E(Y)$ である.』

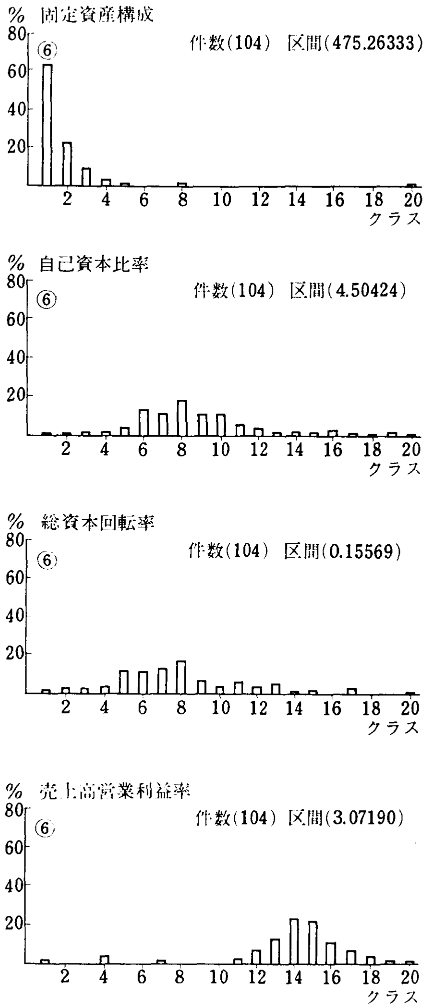
以上の諸仮設によると，$\bar{X}$ 及び $\tilde{X}$ は共に不偏推定量であるから，抽出法選択の基準は，尤度等を慮外すれば，差当り (2) 及び (3) 式の推定精度の比較におくことが出来よう．従って

課題 『仮設 1, 2 の下で仮設 3 の推定諸方式をその分散によって検討し，標本抽出法を決定する事を目的とする.』



第1図 鉄鋼業に関する資本金額1億円以上10億円未満層の規模分布

6) 実際には補足的な意味で，3(D) に於て復元等確率抽出法による比推定法を検討に加えた.



第2図　鉄鋼業に関する資本金額1億円以上10億円未満層の財務諸化率の分布

3. パレート母集団における単純復元抽出法

(A) 母集団の構造

仮設1に従って各種の母集団特性量が得られる．具体的に平均 μ_x, μ_y；分散 $\sigma_x{}^2$, $\sigma_y{}^2$；相関係数 ρ_{xy} について次式が成立する．即ち

(4)
$$\mu_x=\frac{\alpha}{\alpha-1}\theta_1 \qquad \mu_y=\frac{\alpha}{\alpha-1}\theta_2$$

(5)
$$\sigma_x{}^2=\begin{cases}\infty & (1<\alpha\leqq 2)\\ \dfrac{\alpha}{(\alpha-1)^2(\alpha-2)}\theta_1{}^2 & (2<\alpha)\end{cases}$$

$$\sigma_y{}^2=\begin{cases}\infty & (1<\alpha\leqq 2)\\ \dfrac{\alpha}{(\alpha-1)^2(\alpha-2)}\theta_2{}^2 & (2<\alpha)\end{cases}$$

及び $\alpha>2$ に対して

$$\rho_{xy}=\frac{1}{\alpha} \tag{6}$$

である.

(B) 単純復元抽出法の限界と特性

仮設2及び3によると，母平均，推定量 $\bar{X}$ 及び $\tilde{X}$ の分散はそれぞれ

$$\sigma_{\bar{X}}{}^2=\frac{\sigma_x{}^2}{n} \tag{7}$$

$$\sigma_{\tilde{X}}{}^2=\frac{\sigma_x'^2}{n} \tag{8}$$

である．ここで $\sigma_x'^2$ は

$$\sigma_x'^2=\mu_y E\left(\frac{X^2}{Y}\right)-\mu_x{}^2 \tag{9}$$

で表わされる一つの母集団特性量であるが，仮設1に基づいて計算すると部分積分法の反復により容易に

$$\sigma_x'^2=2\,\frac{\alpha^2+1}{(\alpha+1)(\alpha-1)^3}\,\theta_1{}^2 \tag{10}$$

が得られる．この時 C 及び C' を

$$C=\frac{\sigma_x}{\mu_x},\quad C'=\frac{\sigma_x'}{\mu_x} \tag{11}$$

として，$\bar{X}$ 及び $\tilde{X}$ の変動係数を考えると，

$$\text{C.V of }\bar{X}=\frac{c}{\sqrt{n}},\ \text{C.V of }\tilde{X}=\frac{c'}{\sqrt{n}} \tag{12}$$

となるから，標本数を一定にすれば，両式の効果は，C 及び C' により評価される．一方 (4)，(5) 及び (11) 式を考慮すると容易に

$$C=\begin{cases}\infty & (1<\alpha\leqq 2)\\ \dfrac{1}{\sqrt{\alpha(\alpha-2)}} & (2<\alpha)\end{cases} \tag{13}$$

及び

$$C'=\frac{\sqrt{2}}{\alpha}\sqrt{\frac{\alpha^2+1}{\alpha^2-1}}\quad(1<\alpha) \tag{14}$$

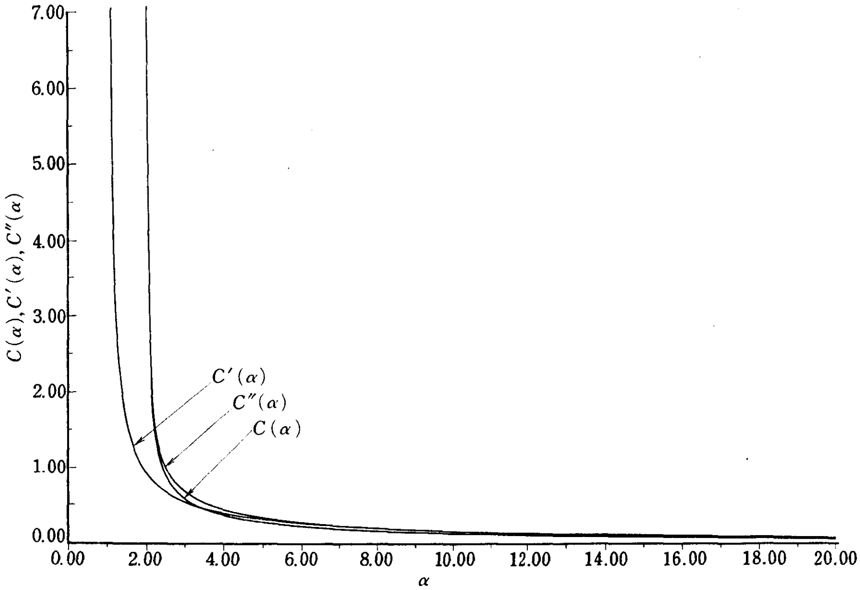
が得られる．つまり，$1<\alpha\leqq 2$ の場合は必然的に (3) の推定方式をとらざるを得ない．従って

結論 1 『各国における多年の測定結果では，所得，資産分布は，

$$1<\alpha\leqq 2 \tag{15}$$

とされるのが普通である．従って方式 (3) は実際上に於ても殆んど絶対的に採用せざるを得ない．』

更に，ここで注目すべきことは，(14) 式及び第3図から明かなように $C'=C'(\alpha)$ は，α の単調減少関数となることである．ところで α が小である事は，一方では，分布 (1) の歪みが大である事を意味し，他方では，(6) 式によって，x, y の相関が大である事を意味する．此等の事実を考えると規模比例確率抽出法に関して，一般に受容されている観点に一見相反する及至



第3図　a に対する変動係数の変化

制限する次のような結論に達するのである.

結論 2. 『規模比例確率抽出法は歪みの大なる分布程有効であると考えられるが, それは等確率抽出法よりも効果が大きいと解すべきであって, 精度そのものは, 比の場合に於ても歪みが大である程低下する.

従って歪みが充分大である場合は, 等確率抽出法における場合と同様に, 全数調査の併用及び層別法, 最適割当法の適用を考慮すべきである.』

結論 3. 『規模比例確率抽出法の補助変量 y は x との相関の大なる程好ましいとされているが, それは正確ではない. その選択は x, y に関する同時密度分布に遡って考察さるべきであり, 例えばそれに適合するパレート係数の大小によって判定さるべきである[7].』

(C) 標本抽出法の精度の比較

(B) に於て, 経験的に容認される所得分布に対しては, 規模比例確率抽出法によるべきであるとする絶対的基準を示した.

然し此の領域を除くと, 常に (4) による推計方式が (3) による推計方式に優先するという根拠はない. 従って, 本節では, 専ら両者の比較可能領域

(16)　$$2 < a$$

に於て, 両方式の検討を試みる.

---

7) (1) 及び (6) 式によれば, パレート係数 a や x, y の相関は x の周辺分布のみによって決定されることになるが, (1) 式は極度に理想化された関係であることに注目すべきである. 例えば (1) 式の代りに

$$P(X \geq x,\ Y \geq y) = \left(\frac{\theta_1{}^p \theta^q{}_2}{\theta_2{}^q x^p + \theta_1{}^p y^q - \theta_1{}^p \theta_2{}^q} \right)^a$$

$$0 < \theta_1 \leq x \leq \infty,\ 0 < \theta_2 \leq y \leq \infty$$

$$1 < a,\ 0 < m,\ n$$

を考察すれば, x の周辺分布はパレート係数 pa をもつパレート分布であるが, p, q, a の値や ρ_{xy} は, ma を固定しても補助変量 y のとり方に影響される事が理解出来よう.

今 R^2 を

$$R^2=\frac{{\sigma_x{}'}^2}{{\sigma_x}^2} \tag{17}$$

で定義すると，(5) 式及び (10) 式により容易に

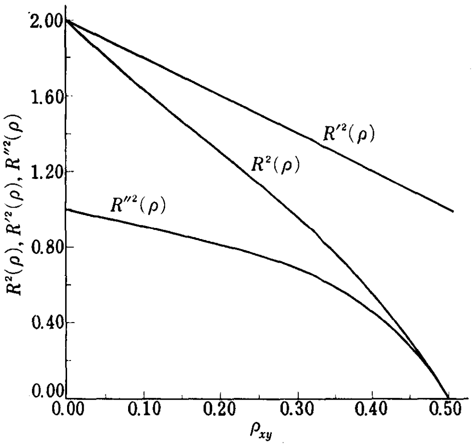
$$R^2=\begin{cases}2\dfrac{(\alpha-2)(\alpha^2+1)}{\alpha(\alpha-1)(\alpha+1)}=2\left(1+\dfrac{1}{\alpha}-\dfrac{2}{\alpha+1}\right)\left(1-\dfrac{1}{\alpha-1}\right)<2 & (2<\alpha),\\ 0 & (1<\alpha<2)\end{cases} \tag{18}$$

或いは (6) 式により

$$R^2=\begin{cases}R^2(\rho)=2(1-2\rho_{xy})\dfrac{1+{\rho_{xy}}^2}{1-{\rho_{xy}}^2} & (0<\rho_{xy}<0.5)\\ 0 & (0.5\leqq\rho_{xy})\end{cases} \tag{19}$$

が得られる.

従って R^2 は (19) 式及び第4図から明かなように ρ_{xy} の単調減少関数である.



第4図 相関に対する分散比の変化

且つ $\alpha=3$ のとき $R^2=\sqrt{\dfrac{5}{6}}<1$

$\alpha=4$ のとき $R^2=\sqrt{\dfrac{17}{15}}>1$

であるから，$3<\alpha<4$ 或いは $\dfrac{1}{4}<\rho_{xy}<\dfrac{1}{3}$ の一点に於て，σ'^2_x と σ^2_x はその大小関係を入れ替えることになる．このことから次のような結論を得ることが出来る.

結論 4. 『規模比例確率抽出法は，歪みが大なる分布に対する程，等確率抽出法に対して相対的に有利である．これは結論2を補完する命題である.』

更に，相関係数 ρ_{xy} を用いてより具体的に表現すると，

結論 5. 『規模比例確率抽出法による推定 (3) は，x, y 間の相関がある程度以上に高い時は，等確率抽出法による推定 (2) よりも精度が高い．より正確にいうと

$\rho_{xy}\geqq 0.5$ ならば，絶対的に規模比例確率抽出法によるべきであり，

$0.5>\rho_{xy}\geqq 0.3$ ならば，規模比例確率抽出法が，相対的に有利であり

$\rho_{xy}<0.3$ ならば，寧ろ等確率抽出法による方が有利である．但し，その精度の向上は，分散比に於て2を超える事はない.』

(D) 比推定法についての検討

以上の各節で得られた結論は，専ら復元等確率抽出の場合は，標本平均により推定することを前提としたものであった．今此の抽出法の下で x に関する母平均を次の比推定方式によった場合を考察してみよう．すなわち，

$$\hat{X}=\mu_y\frac{\bar{X}}{\bar{Y}} \tag{20}$$

或いは

(21) $$\hat{X}'=\mu_y\left(\frac{1}{n}\Sigma\frac{X}{Y}\right)$$

を用いることにする[8]．

此等の推定方式による分散は，周知のように近似的に

(22) $$\sigma_{\hat{X}}^2\sim\sigma_{\hat{X}'}^2\sim\frac{\mu_x^2}{n}\left\{\frac{\sigma_x^2}{\mu_x^2}+\frac{\sigma_x^2}{\mu_y^2}-2\frac{\mathrm{cov}(x,\ y)}{\mu_x\mu_y}\right\}$$

である．従って (1) 式で表われるパレート分布の場合には，容易に (4)，(5)，(6) 及び (13) 式により

(23) $$\mathrm{C.V\ of\ }\hat{X}\sim\mathrm{C.V\ of\ }\hat{X}'\sim\frac{1}{\sqrt{n}}\frac{1}{\alpha}\sqrt{\frac{2(\alpha-1)}{\alpha-2}}$$

となる．(22) 式により明かなように此の方式の適用条件も又 $\alpha>2$ であるが，特に

(24) $$C''=\frac{1}{\alpha}\sqrt{\frac{2(\alpha-1)}{\alpha-2}}\quad\alpha>2$$

とした場合，C'' のグラフは第 3 図に示す通りである．

扨，$\alpha>2$ の場合に R^2 に準じて

(25) $$R'^2=\frac{\sigma_{\hat{X}}^2}{\sigma_{\bar{X}}^2}\sim\left(\frac{C''}{C}\right)^2$$ [9]

を構成すると，R'^2 は推定方式 (2) と (20) 或いは (21) に関する選択基準を与えることになる．

此の場合，(12)，(13) 及び (24) 式を考慮すると

(26) $$R'^2=\frac{2(\alpha-1)}{\alpha}=2\left(1-\frac{1}{\alpha}\right)\quad(\alpha>2)$$

と算定されるから常に

(27) $$1<R'^2<2\quad(\alpha>2)$$

が成立する．一方 R'^2 は $\rho_{xy}=\frac{1}{\alpha}$ の単調減少関数であるから，ρ_{xy} が充分大であれば比推定方式は標本平均方式と大差はないが，一般には，不利な推定方式といえる．つまり

結論 6. 「(1) 式で示されるような歪み型分布に対して，復元等確率抽出法による比推定は標本平均による推定に対して精度上有効であるとはいえない．比推定法を標本平均による推定法より有利とする条件は周知のように $\rho_{xy}>\frac{1}{2}\frac{Y\text{の C.V}}{X\text{の C.V}}$ 或は x の y に関する回帰係数 $\rho_{xy}\frac{\sigma_x}{\sigma_y}$ が平均の比 $\frac{\mu_x}{\mu_y}$ の $\frac{1}{2}$ より大であって歪みの伴わない分布に対して効果をもつものと思われる．」

4. パレート母集団における層化復元抽出法

(A) 層の構成

前節 3(A) における結論 2 によると，規模比例抽出法に於ても猶目層別法が必要であった．

層の構成については，又多くの問題が所在するが，法人企業統計を初めとして，(1) 式で近似されるような歪みの強い分布に対して y を基準にした階層分点 y_ν は，現行に於て屢々 (28)

8) Hartley-Ross の不偏比推定方式による場合も略 (20) 及び (21) 式による場合と同様な結果が得られるであろう．

9) 一方，$R''^2=\frac{\sigma_{\bar{X}}^2}{\sigma_{\hat{X}}^2}\sim\left(\frac{C'}{C''}\right)^2$ を構成すると $R''^2\sim\frac{(\alpha-2)(\alpha^2+1)}{(\alpha+1)(\alpha-1)^2}<1$ が成立する．

式に示すような指数法則に近い形で与えられている．すなわち，この場合

(28) $$y_\nu = a^\nu \theta_2 \quad \nu = 0,\ 1,\ 2,\ \cdots$$

とすることが出来よう．

扨，このような層別を前提とした場合，第 k 層

(29) $$a^{k-1}\theta_2 \leqq y < a^k \theta_2 \quad ; k \geqq 1$$

についての密度分布は

(30) $$f(x,\ y) = A \frac{\alpha(\alpha+1)\theta_1^{\alpha+1}\theta_2^{\alpha+1}}{(\theta_2 x + \theta_1 y - \theta_1\theta_2)^{\alpha+2}} \quad ; 1 < a,$$

$$\theta_1 \leqq x \leqq \infty,\quad a^{k-1}\theta_2 \leqq y < a^k\theta_2$$

(31) $$A^{-1} = a^{-(k-1)\alpha} - a^{-k\alpha}$$

と表わされるから，3.(A) に準じて諸特性量を計算すると，

(32) $$\mu_x^{(k)} = E(X) = \left\{1 + \frac{1}{\alpha-1}\frac{a^{-(k-1)(\alpha-1)} - a^{-k(\alpha-1)}}{a^{-(k-1)\alpha} - a^{-k\alpha}}\right\}\theta_1$$

(33) $$\mu_y^{(k)} = E(Y) = \frac{\alpha}{\alpha-1}\frac{a^{-(k-1)(\alpha-1)} - a^{-k(\alpha-1)}}{a^{-(k-1)\alpha} - a^{-k\alpha}}\theta_2$$

(34) $$E(X^2) = \begin{cases} \left\{1 + \dfrac{2}{\alpha-1}\dfrac{a^{-(k-1)(\alpha-1)} - a^{-k(\alpha-1)}}{a^{-(k-1)\alpha} - a^{-k\alpha}} + \dfrac{2}{(\alpha-1)(\alpha-2)} \right. \\ \qquad \left. \dfrac{a^{-(k-1)(\alpha-2)} - a^{-k(\alpha-2)}}{a^{-(k-1)\alpha} - a^{-k\alpha}}\right\}\theta_1^2 \quad (\alpha \neq 2) \\ \left\{1 + 2\dfrac{a^{-(k-1)} - a^{-k}}{a^{-2(k-1)} - a^{-2k}} + 2\log a\right\}\theta_1^2 \quad (\alpha = 2) \end{cases}$$

(35) $$E(XY) = \begin{cases} \left[\dfrac{\alpha}{\alpha-1}\{a^{-(k-1)(\alpha-1)} - a^{-k(\alpha-1)}\} + \dfrac{1}{\alpha-2}\{a^{-(k-1)(\alpha-2)} - a^{-k(\alpha-2)}\}\right] \\ \dfrac{\theta_1\theta_2}{a^{-(k-1)\alpha} - a^{-k\alpha}} \quad (\alpha \neq 2) \\ [2\{a^{-(k-1)} - a^{-k}\} + \log a]\dfrac{\theta_1\theta_2}{a^{-2(k-1)} - a^{-2k}} \quad (\alpha = 2) \end{cases}$$

(36) $$E(Y^2) = \begin{cases} \dfrac{\alpha}{\alpha-2}\dfrac{a^{-(k-1)(\alpha-2)} - a^{-k(\alpha-2)}}{a^{-(k-1)\alpha} - a^{-k\alpha}}\theta_2^2 \quad (\alpha \neq 2) \\ \dfrac{2\log a\,\theta_2^2}{a^{-2(k-1)} - a^{-2k}} \quad (\alpha \neq 2) \end{cases}$$

及び

(37) $$E\left(\frac{X^2}{Y}\right) = \left\{\frac{\alpha}{\alpha+1}\frac{a^{-(k-1)(\alpha+1)} - a^{-k(\alpha+1)}}{a^{-(k-1)\alpha} - a^{-k\alpha}} + \frac{2}{\alpha} + \frac{2}{(\alpha-1)^2}\right.$$
$$\left.\frac{a^{-(k-1)(\alpha-1)} - a^{-k(\alpha-1)}}{a^{-(k-1)\alpha} - a^{-k\alpha}}\right\}\frac{\theta_1^2}{\theta_2}$$

等が得られる．

(B) 諸特性量の近似

精度の検討を行う場合，(A) の結果をそのまゝ利用することは可成り困難である．然し，今

実際に則して $a=10$ とすれば

(38) $$a \geq 3 \qquad k \geqq 3$$

に於ては次の近似によって大差は生じないであろう．すなわち

(39) $$\mu_x^{(k)} \sim \left\{1+\frac{1}{a-1}\,\frac{a^{(k-1)\alpha}}{a^{(k-1)(\alpha-1)}}\right\}\theta_1 \sim \frac{a^{k-1}}{a-1}\,\theta_1$$

(40) $$\mu_y^{(k)} \sim \frac{a}{a-1}\,a^{k-1}\,\theta_2$$

及び

(41) $$\sigma_x^{(k)2}=E(X^2)-\mu_x^{(k)2} \sim \left\{1+\frac{2}{a-1}\,a^{k-1}+\frac{2}{(a-1)(a-2)}\,a^{2(k-1)}\right\}\theta_1^2$$
$$-\left(1+\frac{a^{k-1}}{a-1}\right)^2\theta_1^2=\frac{a}{(a-1)^2\,(a-2)}\,a^{2(k-1)}\,\theta_1^2$$

(42) $$E(X-\mu_x^{(k)})\,(Y-\mu_y^{(k)})=E(XY)-\mu_x^{(k)}\,\mu_y^{(k)}$$
$$\sim\left\{\frac{a}{a-1}\,a^{k-1}+\frac{a^{2(k-1)}}{a-2}\right\}\theta_1\,\theta_2-\frac{a}{a-1}\,a^{k-1}\left(1+\frac{a^{k-1}}{a-1}\right)\theta_1\,\theta_2$$
$$=\frac{1}{(a-1)^2\,(a-2)}\,a^{2(k-1)}\,\theta_1\,\theta_2$$

(43) $$\sigma_y^{(k)2}=E(Y^2)-\mu_y^{(k)2}$$
$$\sim\left\{\frac{a}{a-2}-\left(\frac{a}{a-1}\right)^2\right\}a^{2(k-1)}\quad\theta_2^2=\frac{a}{(a-1)^2\,(a-2)}\,a^{2(k-1)}\,\theta_2^2$$

従って

(44) $$\rho_{xy}^{(k)}=\frac{E(X-\mu_x^{(k)})\,(Y-\mu_y^{(k)})}{\sigma_x^{(k)}\,\sigma_y^{(k)}} \sim \frac{1}{a}$$

が得られる．一方

(45) $$a \geqq 2 \qquad k \geqq 3$$

に於て

(46) $$\sigma_x^{(k)\prime 2}=\mu_y^{(k)}\,E\left(\frac{X^2}{Y}\right)-\mu_x^{(k)2} \sim \left\{\frac{a}{a+1}\,a^{-(k-1)}+\frac{2}{a}+\frac{2}{(a-1)^2}\,a^{k-1}\right\}\frac{\theta_1^2}{\theta_2}$$
$$\times\frac{a}{a-1}\,a^{k-1}\,\theta_2-\left(1+\frac{a^{k-1}}{a-1}\right)^2\theta_1^2$$
$$=\frac{a}{a-1}\left\{\frac{a}{a+1}+\frac{2}{a}\,a^{k-1}+\frac{2}{(a-1)^2}\,a^{2(k-1)}\right\}\theta_1^2-\left(1+\frac{a^{k-1}}{a-1}\right)^2\theta_1^2$$
$$\sim\frac{a+1}{(a-1)^3}\,a^{2(k-1)}\,\theta_1^2$$

が成立し，更に (38) の条件下で

(47) $$\mu_x^{(k)2}\left\{\frac{\sigma_x^{(k)2}}{\mu_x^{(k)2}}+\frac{\sigma_y^{(k)2}}{\mu_y^{(k)2}}-2\,\frac{\mathrm{cov}^{(k)}(x,\,y)}{\mu_x^{(k)}\,\mu_y^{(k)}}\right\}$$

$$\sim\left(1+\frac{a^{k-1}}{a-1}\right)^2\theta_1^2\left\{\frac{\dfrac{a}{(a-1)^2(a-2)}a^{2(k-1)}\theta_1^2}{\left(1+\dfrac{a^{k-1}}{a-1}\right)^2\theta_1^2}\right.$$

$$\left.+\frac{\dfrac{a}{(a-1)^2(a-2)}a^{2(k-1)}\theta_2^2}{\left(\dfrac{a}{a-1}\right)^2a^{2(k-1)}\theta_2^2}-2\frac{\dfrac{1}{(a-1)^2(a-2)}a^{2(k-1)}\theta_1\theta_2}{\left(1+\dfrac{a^{k-1}}{a-1}\right)\dfrac{a}{a-1}a^{(k-1)}\theta_1\theta_2}\right\}$$

$$\sim\frac{1}{(a-1)^2(a-2)}\left(a+\frac{1}{a}-\frac{2}{a}\right)a^{2(k-1)}\theta_1^2$$

$$\sim\frac{a+a}{(a-1)(a-2)}a^{2(k-1)}\theta_1^2.$$

となる.

(C) 標本抽出法に対する層別の効果

本項では a 及び k に対する (38) 或は (45) の制約の下で (13) の近似を用いて，第3節 (B) (C) (D) に準じた考察を加えることにする．その場合

(48) $$C^{(k)}=\frac{\sigma_x{}^{(k)}}{\mu_x{}^{(k)}}\sim\sqrt{\frac{a}{a-2}}>1\quad(a\geq 3\quad k\geqq 3)$$

(49) $$C^{(k)\prime}=\frac{\sigma_x{}^{(k)\prime}}{\mu_x{}^{(k)}}\sim\sqrt{\frac{a+1}{a-1}}>1\quad(a\geqq 2\quad k\geqq 3)$$

及び

(50) $$C^{(k)\prime\prime}=\sqrt{\frac{\sigma_x{}^{(k)2}}{\mu_x{}^{(k)2}}+\frac{\sigma_y{}^{(k)2}}{\mu_x{}^{(k)2}}-2\frac{\mathrm{cov}^{(k)}(x,y)}{\mu_x{}^{(k)}\mu_y{}^{(k)}}}\sim\sqrt{\frac{a^2-1}{a(a+2)}}>1\quad(a\geqq 3,\ k\geqq 3)$$

によって $C^{(k)}$, $C^{(k)\prime}$ 及び $C^{(k)\prime\prime}$ を定義すれば，前節 (B) の所論に従って，此等はそれぞれの抽出法による標本精度の基準を与えることになる．

注目すべき事は，此等の層に於ては何れにしても $C^{(k)}$ 等の変動係数が1より小とならないことである．他方 $C^{(k)}$, $C^{(k)\prime}$ 及び $C^{(k)\prime\prime}$ の順位は，次の諸式，すなわち

(51) $$R^{(k)2}=\left\{\frac{C^{(k)\prime}}{C^{(k)}}\right\}^2\sim\frac{(a+1)(a-2)}{a(a-1)}=1-\frac{2}{a(a-1)}<1\quad(a\geqq 3,\ k\geqq 3),$$

(52) $$R^{(k)\prime 2}=\left\{\frac{C^{(k)\prime\prime}}{C^{(k)}}\right\}^2\sim\frac{a^2-1}{a^2}<1\quad(a\geqq 3,\ k\geqq 3)$$

及び

(53) $$R^{(k)\prime\prime 2}=\left\{\frac{C^{(k)\prime}}{C^{(k)\prime\prime}}\right\}^2\sim\frac{a(a-2)}{(a-1)^2}<1\quad(a\geqq 3,\ k\geqq 3)$$

によって直接決定することが出来る．

つまり以上の諸関係から

(54) $$C^{(k)\prime}<C^{(k)\prime\prime}<C^{(k)}\quad(a\geqq 3,\ k\geqq 3)$$

が得られる結果，規模比例確率抽出法の優位性が歴然と示されるのである．

ところで前節 (C) によれば $\alpha \geqq 3$ では一般的には

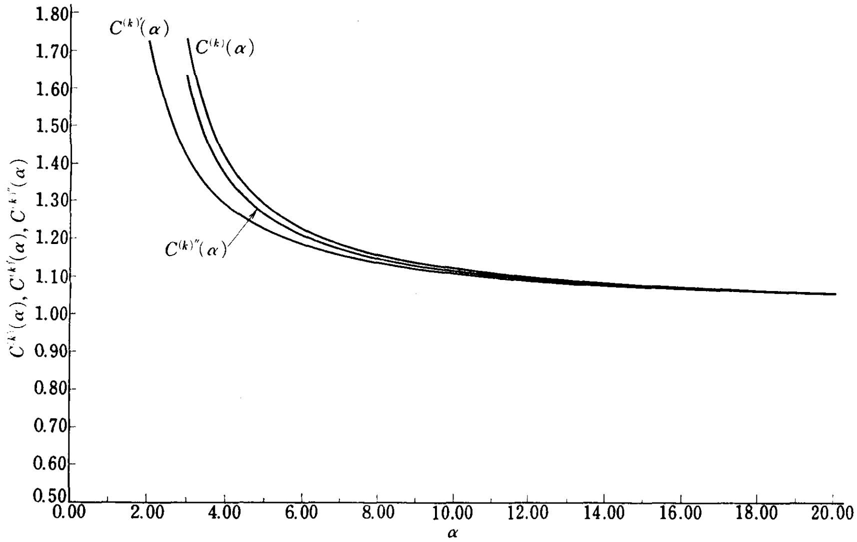
$$C' > C \tag{55}$$

であり，又前節 (D) によれば常に

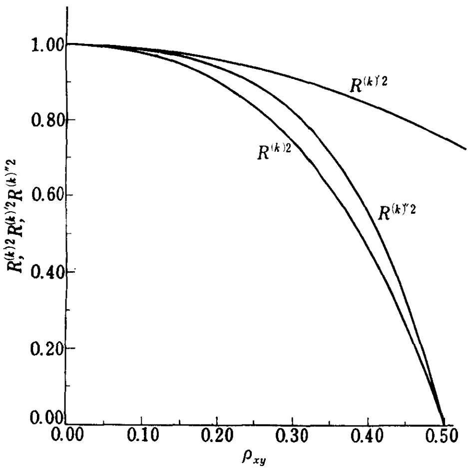
$$C'' > C \tag{56}$$

であった．従って (54) 式は更に (28) の如き層別の効果が特に規模比例確率抽出法及び比推定方式に対して顕著であることを示しているのである．従って本節における結論として，

結論 7. 「第 (1) 式の分布に従う母集図に対して (28) 式で示された層別を加えると，$\alpha \geqq 3$ の場合，$k \geqq 3$ の各層に於ては，一般には $\rho_{xy} \leqq \frac{1}{3}$ であるにも拘らず規模比例確率抽出法が最も有利であり，次いで比推定方式が有効である．層別の効果も此の両者に於て最も著しい．然し乍ら其等の何れにしても $C^{(k)}$ 等の値は1より小さくはならない．」



第5図 層内における変動係数の変化



第6図 層内における分散比の変化

処で以上に於て特に注意さるべきことは，結論7はあくまで (13) の近似が許される範囲に限られることである．従って小規模階層については一般に明確な基準が簡単には得られないばかりでなく，最も現実性の強い $1<\alpha<3$ に関しても容易に単純な結論を下し得ないのである．

もっとも以上各節の結果から推して，α の此の範囲に関しても大規模階層に対して規模比例確率抽出法を有利とする見通しを立てることが出来るが，第7節の実証的検討を俟つことが必要であろう．

5. パレート母集団における復元不等確率抽出法による最適抽出確率

これまで扱って来た規模比例確率抽出法は，抽出確率 p.s を

(57)
$$\mathrm{p.s} = \frac{y^{\nu}}{E(Y^{\nu})}$$

とする復元不等確率抽出法え一般化することが出来る．比の場合 μ_x の不偏推定量 $\tilde{X}^{[\nu]}$ 及びその分散 $\sigma_{\tilde{X}[\nu]}{}^2$ は，それぞれ $\nu<\alpha$ 及び $\alpha+\nu>2$ で存在し

(58)
$$\tilde{X}^{[\nu]} = \mu_y{}^{[\nu]} \sum_{i=1}^{n} \frac{X_i}{Y_i{}^{\nu}}$$

及び

(59)
$$\sigma_{\tilde{X}[\nu]} = \frac{1}{n}\,\sigma_x{}^{[\nu]2}$$

で与えられる．但し

(60)
$$\mu_\nu{}^{(\nu)} = E(Y^{\nu})$$

及び

(61)
$$\sigma_x{}^{[\nu]2} = \mu_y{}^{(\nu)}\,E\left(\frac{X^2}{Y^{\nu}}\right) - \mu_x{}^2$$

である．従って今

(62)
$$\alpha > \nu,\quad \alpha + \nu > 2$$

とすると $\sigma_x{}^{[\nu]2}$ は

(63)
$$\sigma_x{}^{[\nu]2} = \left\{\frac{\alpha}{\alpha+\nu} + \frac{2}{\alpha+\nu-1} + \frac{2}{(\alpha-1)(\alpha+\nu-2)}\right\}\frac{\alpha}{\alpha-\nu}\,\theta_1{}^2$$

となる．此の際第3節に従って

(64)
$$C^{[\nu]} = \frac{\sigma_x{}^{[\nu]}}{\mu_x}$$

を定義すると容易に (62) の条件の下で

(65)
$$C^{[\nu]} = \sqrt{\frac{\alpha-1}{\alpha-\nu}\left\{\frac{\alpha-1}{\alpha+\nu} + \frac{2(\alpha-1)}{\alpha(\alpha+\nu-1)} + \frac{2}{\alpha(\alpha+\nu-2)}\right\} - 1}$$

が得られる．第7図は ν の種々の値に対する $C^{[\nu]}$ の変化を示すものである．従って最適抽出確率は

(66)
$$\frac{dC^{[\nu]}}{d\nu} = 0$$

によって与えられるが，その結果容易に

$$\frac{1}{\alpha-\nu}\left\{\frac{\alpha-1}{\alpha+\nu}+\frac{2(\alpha-1)}{\alpha(\alpha+\nu-1)}+\frac{2}{\alpha(\alpha+\nu-2)}\right\}-\left\{\frac{\alpha-1}{(\alpha+\nu)^2}+\frac{2(\alpha-1)}{\alpha(\alpha+\nu-1)^2}+\frac{2}{\alpha(\alpha+\nu-2)^2}\right\}=0 \tag{67}$$

が成立する．(67) 式をみたす $\nu^{(0)}=\nu^{(0)}(\alpha)$ は

$$-\alpha<\nu^{(0)}<\alpha \tag{68}$$

の範囲内にあるが，その場合 $C^{[\nu^{(0)}]}=C^{[\nu^{(0)}]}(\alpha)$ は第8図のように (65) 式をみたす曲線群の包絡曲線となる．これによって又標本数を固定した場合，最適抽出確率を与えても猶且 α が少なる範囲では一定限度以上に精度を向上させる得ない事が明かである．逆に今目標精度 e を与えた場合の絶対最少標本数を n_L とすれば，それは

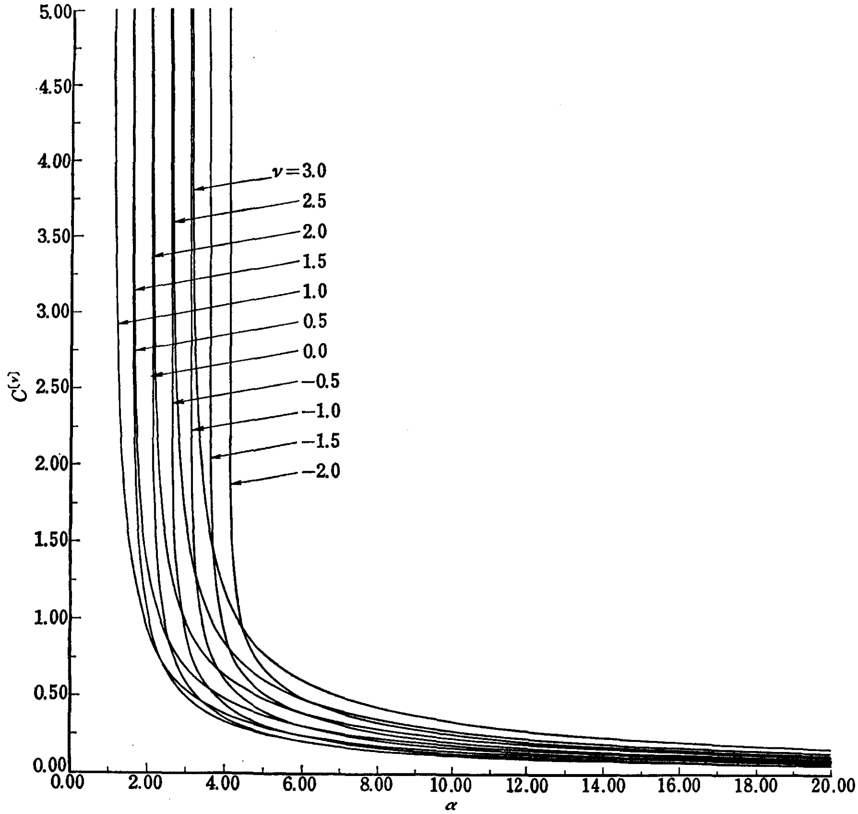
$$\frac{C^{[\nu^{(0)}]}}{n_L}=e \tag{69}$$

によって決定される．

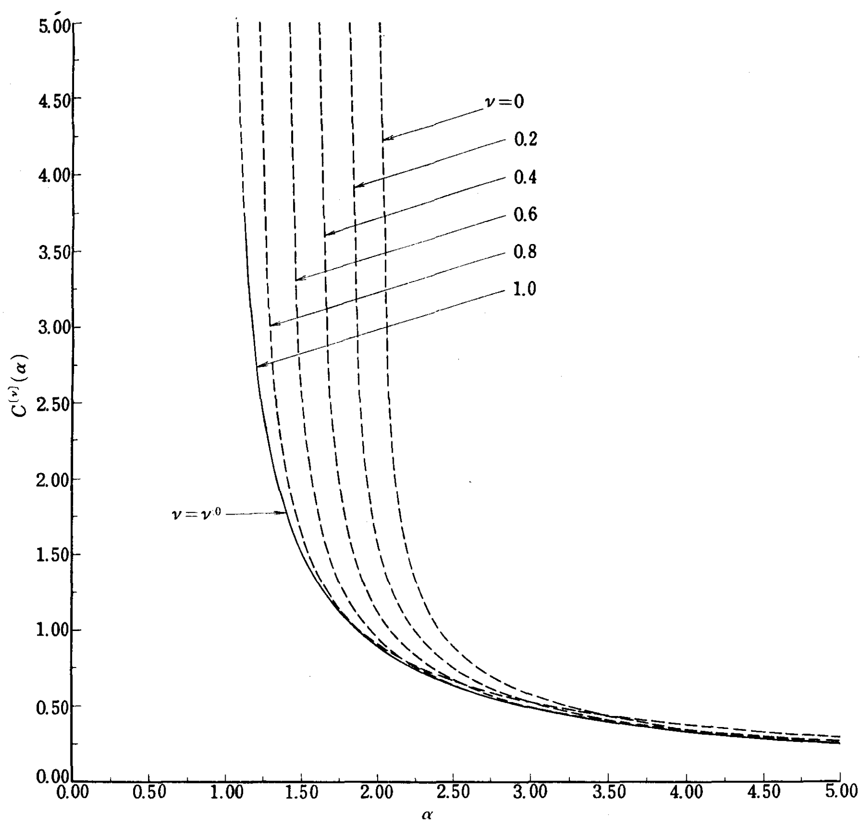
一方，第9図は (67) 式からシミュレーションによって得られた α に対する $\nu^{[0]}$ の曲線であるが，それによると

$$0<\nu^{(0)}<1 \tag{70}$$

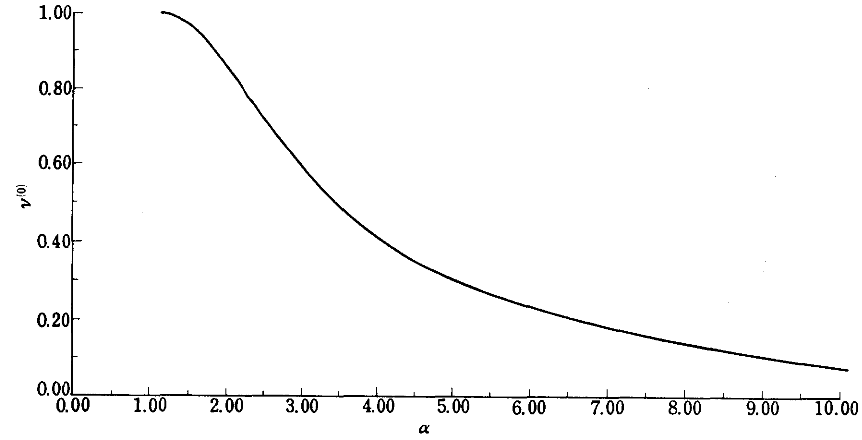
と考える事が出来よう．又 $\nu^{(0)}$ は α に対して単調に減少する傾向をもち，少くとも $\alpha<3.3$



第7図　抽出確率によるCの変動



第8図



第9図　最適抽出比の決定曲線

$\rho_{xy}>3$ に於ては，$\nu^{(0)}$ は0よりは1に近いことになる．従って此の範囲では規模比例確率抽出法が有効であるとする第3節 (C) の結論5に合致する．以上を纏めると

結論 8. 『(57) 式の形で示される不等確率抽出法における最適抽出確率 $\nu^{(0)}$ は (67) 式により決定される．$\nu^{(0)}$ の存在範囲は，シミュレーションによれば第1表に示すように (70) 式に

第1表　α に対する最適抽出比の変動 ($\nu^{(0)}$)

α	$\nu^{(0)}$	α	$\nu^{(0)}$
1.10	0.99954	3.10	0.56602
1.20	0.99632	3.20	0.54425
1.30	0.98956	3.30	0.52438
1.40	0.97866	3.40	0.50505
1.50	0.96373	3.50	0.48626
1.60	0.94583	3.60	0.46923
1.70	0.92480	3.70	0.45268
1.80	0.90163	3.80	0.43694
1.90	0.87657	3.90	0.42211
2.00	0.85036	4.00	0.40765
2.10	0.82326	4.10	0.39466
2.20	0.79569	4.20	0.38164
2.30	0.76794	4.30	0.36963
2.40	0.74045	4.40	0.35824
2.50	0.71327	4.50	0.34776
2.60	0.68658	4.60	0.33718
2.70	0.66066	4.70	0.32773
2.80	0.63560	4.80	0.31812
2.90	0.61137	4.90	0.30920
3.00	0.58810	5.00	0.30068

(以上は Newton 法による (67) の関係の シミュレーションである．但し初期値は $\alpha\leqq 2$ に対して，1，$\alpha>2$ に対して 0 を用いた)

よって与えられる．又 $\nu^{(0)}$ は α の増大と共に単調に減少する．他方 $\nu^{(0)}$ に対応する $C^{[\nu^{(0)}]}$ は α に対して単調に減少するが，α が小なる時，比較的大きな値をとる結果，全数抽出或いは層別等の配慮が必要である.」

6. ジブラ母集団における抽出効果

第2節の当初に述べた経験的な理由によりこれまで専らパレート母集団を対象として，抽出法を検討した．然し乍ら，ジブラ分布は経済統計や生物統計に於て猶重要な地位を占めるものであり，事実特定の現象或いは国に於て極めて高い適合度が示されている．(例えば Aitchison and Brown [1] 参照)．従って此の分布に対して，改めて以上の検討を加えることは極めて有意義であるのみならず，最適抽出確率に関する結果には，極めて興味深いものがある．

(A) 母集団の特性

周知のように，ジブラ分布は，対数正規性をもつものとして，二次元の場合

$$(71)\quad f(x, y)=\frac{1}{2\pi D_x D_y \sqrt{1-r^2}\,xy}\exp\left[-\frac{1}{2(1-r^2)}\right.$$
$$\left.\left\{\frac{(\log x-m_x)^2}{D_x^2}-\frac{2r(\log x-m_y)(\log y-m_y)}{D_x D_y}+\frac{(\log y-m_y)^2}{D_y^2}\right\}\right]$$

によって一般形とすることが出来る．

その際 m_x, m_y; D_x, D_y 及び r は，それぞれ対数平均，ジブラ係数及び対数相関係数を表わしていることは，いうまでもない．

此の分布に対して，平均 μ_x, μ_y, 分散 σ_x^2, σ_y^2 及び相関係数 ρ_{xy} を算出すると，それぞれ

(72) $$\mu_x = e^{\frac{D_x^2}{2}+m_x},\quad \mu_y = e^{\frac{D_y^2}{2}+m_y}$$

(73) $$\sigma_x^2 = e^{2m_x+D_x^2}(e^{D_x^2}-1)$$
$$\sigma_y^2 = e^{m_y+D_y^2}(e^{D_y^2}-1)$$

(74) $$E(X-\mu_x)(Y-\mu_y) = e^{m_x+m_y+\frac{D_x^2+D_y^2}{2}}(e^{rD_xD_y}-1)$$

及び

(75) $$\rho_{xy} = \frac{e^{rD_xD_y}-1}{\sqrt{e^{D_x^2}-1}\sqrt{e^{D_y^2}-1}}$$

が得られる．従って又，x に関する母変動係数 C は

(76) $$C = \sqrt{e^{D_x^2}-1}$$

によって表わされる．

(B) 諸抽出法による精度

標本数を固定した場合，等確率抽出によって得られた標本平均を用いて母平均を推定する場合，その変動係数は勿論 C によって決定される．又，等確率抽出によって得られた標本によって μ_x の比推定を行う場合は

(77) $$C'' = \sqrt{\frac{\sigma_x^2}{\mu_x^2}+\frac{\sigma_y^2}{\mu_y^2}-2\rho_{xy}\frac{\sigma_x\sigma_y}{\mu_x\mu_y}} = \sqrt{e^{D_x^2}-1+e^{D_y^2}-1-2(e^{rD_xD_y}-1)}$$
$$= \sqrt{e^{D_x^2}+e^{D_y^2}-2e^{rD_xD_y}}$$

によって決定されると看似すことが出来る．

ジブラ母集団に於ては，此の両者の選択に対して特に極立った特徴は認められない．

従って問題は y に比例した確率による抽出法に於て生ずるものと思われる．

今 $E\left(\frac{X^2}{Y}\right)$ を算定する為に変換

(78) $$u = \frac{\log x - m_x}{D_x},\quad v = \frac{\log y - m_y}{D_y}$$

を行うと容易に

(79) $$E\left(\frac{X^2}{Y}\right) = \frac{1}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-r^2)}\{u^2-2ruv+v^2\}\right.$$
$$\left.+2m_x+2D_xu-m_y-\frac{D_yv}{2}\right]du\,dv$$
$$= \frac{1}{2\pi\sqrt{1-r^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2(1-r^2)}\right.$$
$$\left.\{(u-p)^2-2r(u-p)(v-q)+v-q)^2\}+w\right]du\,dv = e^w$$

が得られる．但しこゝで p, q, w は

(80) $$\begin{cases}-2p+2rq = 4(1-r^2)D_x\\ 2rp-2q = -2(1-r^2)D_y\end{cases}$$

及び

$$w = 2m_x - D_y^2 + \frac{p^2 - 2rpq + q^2}{2(1-r^2)} \tag{81}$$

の解として，それぞれ

$$\begin{cases} p = -2D_x + rD_y \\ q = D_y - 2r\,D_x \end{cases} \tag{82}$$

及び

$$w = 2m_x - m_y + 2D_x^2 - 2r\,D_x\,D_y + \frac{D_y^2}{2} \tag{83}$$

によって与えられる値である．故に結局

$$E\left(\frac{X^2}{Y}\right) = e^{2m_x - m_y + 2D_x^2 - 2r\,D_x\,D_y + \frac{D_y^2}{2}} \tag{84}$$

が成立する．今 (72) 式を考慮すると

$$\begin{aligned} \sigma_x'^2 &= \mu_y\,E\left(\frac{X^2}{Y}\right) - \mu_x^2 \\ &= e^{2m_x + 2D^2_x - 2r\,D_x\,D_y + D_y^2} - e^{2m_x + D_x^2} \end{aligned} \tag{85}$$

となるから，$C' = \dfrac{\sigma_x'}{\mu_x}$ は

$$C' = \sqrt{e^{D_x^2 - 2r\,D_x\,D_y + D_y^2} - 1} \tag{86}$$

によって示される．従って又 $R^2 = \dfrac{C'^2}{C^2}$ は

$$R^2 = \frac{e^{D_x^2 - 2r\,D_x\,D_y + D_y^2} - 1}{e^{D_x^2} - 1} \tag{87}$$

となるから，もし規模比例確率抽出法が，等確率抽出による標本平均推定法より有利であるとするならば

$$-2r\,D_x\,D_y + D_y^2 < 0$$

故に

$$r\,\frac{D_x}{D_y} > 0.5 \tag{88}$$

を満さねばならない．すなわち x の y に関する対数回帰係数が 0.5 より大であることが必要であり且，大である程有効となるのである．

扨，問題を更に一般化して y^ν に比例する抽出確率による推定 (58) と，その誤差決定量 (61) を考察しよう．此の場合 (79) 式より (84) 式を得たのと同様な過程により，

$$E(Y^\nu) = e^{\nu\,m_y + \frac{\nu^2 D_y^2}{2}} \tag{89}$$

及び

$$E\left(\frac{X^2}{Y^\nu}\right) = e^{2m_x - \nu\,m_y + 2D_x^2 - 2\nu r D_x D_y + \frac{\nu^2}{2} D_y^2} \tag{90}$$

が得られるから

$$\sigma_x^{[\nu]2} = E(Y^\nu)\,E\left(\frac{X^2}{Y^\nu}\right) - \mu_x^2 \tag{91}$$

$$= e^{2m_x+2D_x^2-2\nu rD_xD_y+\nu D_y^2} - e^{2m_x+D_x^2},$$

従って又

$$C^{[\nu]} = \frac{\sigma_x^{[\nu]}}{\mu_x} = \sqrt{e^{D_x^2-2\nu rD_xD_y+\nu D^2_y}-1} \tag{92}$$

が成立する.

故に最適抽出率を決定する $\nu^{(0)}$ は $\dfrac{dC^{[\nu]}}{d\nu}=0$ の解として

$$\nu^{(0)} = r\frac{D_x}{D_y} \tag{93}$$

によって与えられることになる. つまり, この $\nu^{(0)}$ は, x の y に関する対数回帰係数そのものに他ならない. 更にこの $\nu^{(0)}$ に対して

$$C^{[\nu^{(0)}]} = \sqrt{e^{D_x^2(1-r^2)}-1} \tag{94}$$

が成立するから, 対数相関係数 r に対して

$$C^{[\nu^{(0)}]} \to D_x\sqrt{1-r^2} \quad (r\to 1) \tag{95}$$

となる. 此の右辺は x の y に関する対数回帰直線の残差の標準偏差に等しいことが注目される. 以上の結果を纒めると容易に次の結論に達することが出来よう.

結論 9. 「ジブラ母集団に於て y^ν に比例する確率抽出法は, ν が, x の y に関する対数回帰係数のとき最適となる. 従ってこの係数が0に近い程等確率抽出法が有効であり, 1に近い程規模比例確率抽出法が有効となる.

又, 此の最適抽出法による変動係数は対数相関が1に近づく時, 0に限りなく近づく.

且対数相関が1に近い時, 変動係数は近似的に, x の y に関する対数回帰直線の残差の標準偏差に等しい.」

扨パレート母集団を考察した場合に於ては専ら単純相関係数を基準としたが, その結果は層別の有無や, 規模階層の大小等に影響されて一定した結果を示すことが出来なかった.

然し乍ら本節のジブラ母集団の考察によって我々は次のような推測を行うことが出来る.

結論 10. 「規模の累乗に比例した抽出確率をもつ不等確率抽出法の決定基準は, 単純相関係数や単純回帰係数ではなく, 歪みを考慮した相関係数及び回帰係数, 例えば対数相関係数, 対数回帰係数が有望である. このことは, 比推定法と標本平均推定法との優劣が単純相関係数, 回帰係数によって影響される事実と極めて対照的である.」

7. 標本抽出法の撰択基準に関する実証的検討と問題点

これまでの各節における理論的諸検討結果は, 法人企業の抽出法の選択に寄与するものであるが, それは飽くまで理想化されは理論的母集団に関する推論であった. 現実の分布はパレート近似を比較的有効とするが, 猶項目により可成り不充分な場合を生ずるのである. 従って最終的には既存データによる再検討が必要である. その際第5節で扱った最適抽出法を用いることは, 現実の分布の把握が不充分である現状に於ては, 適切であるとは思われない. 従って本節に於ては, 専ら層化等確率抽出法及び層化規模比例確率抽出法のみを対象とした. 此の場合, 第4節に於て理論的把握が困難であった $1<\alpha<3$ が, データに則した状況と思われるので, 本節に於けるデーター解析の結果は, 独自の内容と意義をもつものと云えよう.

処で, 以下で扱われた諸資料は, 諸般の状況を考慮し, 又各種の相関係数を算定し, 比較検討した結果, 補助変量 y としては資本金を用いるのが妥当と思われたので, 専ら其を基準とし

て作製されたものである.

扨，第2表によると，対資本金相関係数は，各産業，各費目を通じて，資本金階層の上昇に伴って増大する傾向にあると認められる.

他方，第3表によれば，C 及び C' は，各産業，各費目を通じて，資本金階層の上昇に対して停滞（C の場合）乃至減少する（C' の場合）傾向を示している.

此等は，相関係数の規模比例確率抽出法に対する効果に関する通説に一致するものであり，他方第4節で予想したように大規模階層に対して規模比例確率抽出法が有利であるとする結果を与えている．然し反面に於て，以上の結果は相関が高い程精度が低下することを示した第3節の結論2及び大規模階層に於ては相関係数及び変動係数は略一定であるとする第4節の結論に合致しないことに留意せねばならない.

第2表 a）　産業別資本金階級別資本金に対する主要項目の相関係数
（44年度法人企業統計年報）

産業	費目	資本金階層					
		2百万円未満	2百万円以上1千万円未満	1千万円以上5千万円未満	5千万円以上1億円未満	1億円以上10億円未満	10億円以上
農林水産業	総資本	0,5000	0,4425	0,0224	0,3425	0,3027	0,9635
	固定資産	0,4083	0,3784	0,0458	0,2064	0,4257	0,9949
	売上高	0,0000	0,4243	0,0173	0,3533	0,2177	0,9207
鉱業	総資本	0,0900	0,0906	0,7585	0,1536	0,7249	0,7723
	固定資産	0,1643	0,0812	0,6429	0,0224	0,7348	0,7797
	売上高	0,0000	0,0300	0,7788	0,0000	0,2997	0,6690
建設業	総資本	0,0000	0,3385	0,5493	0,0316	0,1936	0,9364
	固定資産	0,0995	0,3715	0,3415	0,3035	0,2602	0,9308
	売上高	0,0794	0,3890	0,6764	0,2538	0,3306	0,9469
食料品製造業	総資本	0,0000	0,3271	0,3660	0,0640	0,5920	0,4187
	固定資産	0,0000	0,2933	0,2800	0,0806	0,5684	0,4556
	売上高	0,1556	0,2548	0,3937	0,0361	0,5568	0,4107
繊維工業	総資本	0,3592	0,6112	0,1720	0,4838	0,4129	0,9388
	固定資産	0,3162	0,5897	0,4181	0,0469	0,4690	0,9086
	売上高	0,1985	0,3985	0,0600	0,4835	0,2300	0,8994
紙・化学工業	総資本	0,1281	0,2412	0,4201	0,0100	0,6520	0,9347
	固定資産	0,0000	0,1507	0,4815	0,0548	0,5897	0,9058
	売上高	0,0000	0,1755	0,4268	0,0548	0,3951	0,9037
金属工業	総資本	0,2722	0,3483	0,7699	0,2627	0,6733	0,9929
	固定資産	0,0000	0,2943	0,5436	0,2532	0,6030	0,9925
	売上高	0,3956	0,4221	0,6920	0,1217	0,4488	0,9714
機械工業	総資本	0,0000	0,3491	0,4979	0,2140	0,6785	0,9419
	固定資産	0,0000	0,3444	0,5579	0,0663	0,6224	0,9208
	売上高	0,0000	0,3000	0,5161	0,2486	0,6058	0,9153
卸売業	総資本	0,1655	0,2124	0,3476	0,3865	0,3077	0,8426
	固定資産	0,0000	0,2608	0,3695	0,4068	0,2466	0,8435
	売上高	0,1127	0,2052	0,3106	0,3606	0,5604	0,8798
小売業	総資本	0,1849	0,2119	0,5404	0,0990	0,2903	0,8883
	固定資産	0,0000	0,1091	0,2209	0,3186	0,4759	0,9489
	売上高	0,3292	0,2241	0,3274	0,3863	0,3429	0,9101
不動産業	総資本	0,1957	0,2729	0,0000	0,0624	0,1691	0,9630
	固定資産	0,1913	0,2929	0,1947	0,0200	0,2236	0,9502
	売上高	0,2086	0,0316	0,0854	0,0735	0,0000	0,9710
運輸通信業	総資本	0,2236	0,4764	0,5865	0,0980	0,7446	0,8236
	固定資産	0,0000	0,4163	0,6219	0,2625	0,7944	0,7687
	売上高	0,5570	0,4399	0,2247	0,3018	0,5866	0,9287
サービス業	総資本	0,2722	0,2280	0,2302	0,0900	0,2902	0,5631
	固定資産	0,0000	0,2737	0,2559	0,2802	0,5913	0,6832
	売上高	0.3375	0,1640	0,0872	0,0346	0,2057	0,6956

第2表 b) 産業別資本金階級別資本金に対する主要項目の相関係数
(45年度法人企業統計季報1〜3月分)

産業	費目	資本金階層				
		2百万円以上 1千万円未満	1千万円以上 5千万円未満	5千万円以上 1億円未満	1億円以上 10億円未満	10億円 以上
農林水産業	総資本	0.1688	0.5137	0.3156	0.4738	0.8964
	固定資産	0.1269	0.4595	0.5568	0.5410	0.9889
	売上高	0.1095	0.2623	0.2170	0.4264	0.7392
鉱業	総資本	0.1054	0.1729	0.0000	0.8028	0.7684
	固定資産	0.0700	0.1296	0.2522	0.8514	0.7857
	売上高	0.1249	0.2383	0.0678	0.6508	0.6785
建設業	総資本	0.3816	0.4603	0.6421	0.7424	0.9222
	固定資産	0.1649	0.0860	0.3752	0.6786	0.9025
	売上高	0.2789	0.3855	0.3341	0.5123	0.9303
食料品製造業	総資本	0.1562	0.7106	0.2919	0.5647	0.7918
	固定資産	0.0800	0.6353	0.5723	0.5688	0.8858
	売上高	0.0975	0.5372	0.4207	0.2717	0.7859
繊維工業	総資本	0.1253	0.1980	0.4494	0.8079	0.9810
	固定資産	0.0424	0.3242	0.1068	0.7436	0.9794
	売上高	0.1200	0.1720	0.4857	0.6994	0.9656
紙・化学工業	総資本	0.0510	0.3658	0.1881	0.6620	0.8863
	固定資産	0.0721	0.3499	0.5744	0.5773	0.8627
	売上高	0.1833	0.2740	0.0173	0.4882	0.8285
金属工業	総資本	0.0346	0.1000	0.6619	0.5711	0.9936
	固定資産	0.0224	0.2193	0.7303	0.6400	0.9908
	売上高	0.0173	0.1533	0.4921	0.3965	0.9771
機械工業	総資本	0.3662	0.1926	0.1670	0.6801	0.9426
	固定資産	0.4946	0.2184	0.1691	0.7090	0.8889
	売上高	0.3890	0.1744	0.2081	0.5426	0.9364
その他製造業	総資本	0.3722	0.1175	0.0245	0.2563	0.8772
	固定資産	0.4271	0.2839	0.3647	0.3017	0.8789
	売上高	0.2612	0.1556	0.0906	0.2022	0.7364
卸売業	総資本	0.3876	0.4253	0.0361	0.5748	0.8912
	固定資産	0.2352	0.3077	0.0361	0.6508	0.9380
	売上高	0.4060	0.4484	0.0224	0.4280	0.9103
小売業	総資本	0.3844	0.6603	0.7261	0.4691	0.9133
	固定資産	0.3259	0.6341	0.4141	0.6243	0.9616
	売上高	0.2126	0.6987	0.2638	0.3393	0.9572
不動産業	総資本	0.0173	0.1811	0.2980	0.2680	0.9130
	固定資産	0.0200	0.0400	0.3023	0.1884	0.9307
	売上高	0.1503	0.3431	0.2627	0.2358	0.6776
運輸通信業	総資本	0.2335	0.4652	0.1109	0.5414	0.9507
	固定資産	0.1497	0.4435	0.1356	0.5986	0.9582
	売上高	0.1229	0.3697	0.1356	0.2000	0.8173
サービス業	総資本	0.5952	0.2551	0.0787	0.3695	0.9138
	固定資産	0.5280	0.2025	0.0735	0.5790	0.8685
	売上高	0.5071	0.1068	0.1204	0.0000	0.8730

処で更に第4表によると，資査金階層の上昇に伴ない，R^2 が小となる顕著な傾向を，産業別及び費目別読みとることが出来るであろう．又 $R^2=1$ を与える対応点を資本金について求めると，此の表から略1億円前後と読み取ることが出来よう．他方，第2表によると，この分点は，同時に $\rho_{xy}=0.3$ に略対応しているようである．

猶又第4表の全体を通して，$R^2<2$ として大きな危険はないようである．従って係数 R^2 に関しては，第3節の結論4及び5が妥当すると共に，第4節における結論7とは必ずしも合致

第3表 項目別の C 及び C' の変化（その1）
（46年法人企業統計年報）

産業	項目	資本金階層 5百万円以上1千万円未満 C	C′	1千万円以上1億円未満 C	C′	1億円以上10億円未満 C	C′	10億円以上50億円未満 C	C′	50億円以上 C	C′
化学工業	売掛金	1,3782	1,0967	1,1225	0,9005	1,1314	1,4275	0,8644	0,6652	1,0176	0,5019
	流動資産	1,0378	0,8567	1,0292	0,8038	0,9581	1,0978	0,7768	0,5903	0,9415	0,3779
	固定資産	0,9198	0,8755	0,9341	0,6892	1,1328	1,3161	0,9774	0,7414	0,9459	0,3794
	資産合計	0,9282	0,7595	0,9333	0,6953	0,9184	1,0696	0,8080	0,5882	0,9235	0,3387
	買掛金	1,6335	1,4342	0,9883	0,9302	1,0820	1,1836	0,9053	0,7245	0,9422	0,3915
	流動負債	1,1142	0,9432	1,0244	0,8492	0,9800	1,0830	0,8367	0,6463	0,9930	0,4229
	固定負債	1,6319	3,1363	1,2139	1,0896	1,3399	1,3701	1,1658	0,9743	0,9385	0,4751
	自己資本	1,7273	1,2172	1,6295	1,3495	1,2709	1,6799	0,9292	0,6813	1,0037	0,4484
	売上高	1,0063	0,8739	0,9831	0,8182	0,9334	1,0742	0,9199	0,7400	1,0145	0,3533
	従業員	0,8228	0,9437	1,0247	0,7505	1,0576	1,3877	0,8913	0,7249	0,9822	0,4473

（その2）

産業	項目	資本金階層 1千万円以上5千万円未満 C	C′	5千万円以上1億円未満 C	C′	1億円以上10億円未満 C	C′	10億円以上 C	C′
鉄鋼	売掛金	1,2281	1,0194	1,2106	1,1431	1,0342	0,9389	2,0131	0,6651
	流動資産	1,6734	1,5202	1,3691	1,4909	1,2428	0,8573	2,4575	0,3902
	固定資産	1,1121	0,9476	0,8981	0,9657	1,1508	0,9434	2,9363	0,3205
	資産合計	1,2785	1,0197	0,8282	0,8264	0,9978	0,7763	2,6430	0,2797
	買掛金	1,7608	1,4131	1,3017	1,2574	1,0382	0,9272	2,0102	0,5880
	流動負債	2,0821	2,6569	1,7202	1,7066	2,2660	1,4888	2,3959	0,5427
	固定負債	1,2850	1,2445	1,3106	1,3279	1,9422	1,5086	3,0299	0,4089
	自己資本	1,3399	1,1953	1,0687	1,0678	1,2525	1,1901	2,2310	0,6923
	売上高	1,1451	1,0640	0,9735	0,9908	0,8885	0,7925	2,4293	0,4435
	従業員	0,7419	0,8194	0,6159	0,6269	1,0830	1,2916	2,3421	0,4450

第4表 a) 項目別の R^2 の変化（その1）
（46年法人企業統計年報）

産業	項目	資本金階層				
		5百万円以上 1千万円未満	1千万円以上 1億円未満	1億円以上 10億円未満	10億円以上 50億円未満	50億円 以上
化学工業	売掛金	0,6332	0,6436	1,5919	0,5922	0,2432
	流動資産	0,6815	0,6099	1,3129	0,5775	0,1611
	固定資産	0,9061	0,5443	1,3498	0,5775	0,1608
	資産合計	0,6695	0,5550	1,3563	0,5299	0,1345
	買掛金	0,7709	0,8858	1,1966	0,6404	0,1727
	流動負債	0,7167	0,6872	1,2213	0,5967	0,1814
	固定負債	3,6938	0,8057	1,0456	0,6984	0,2563
	自己資本	0,4966	0,6858	1,7472	0,5375	0,1996
	売上高	0,7542	0,6926	1,3244	0,6472	0,1213
	従業員	1,3155	0,5364	1,7217	0,6615	0,2074

（その2）

産業	項目	資本金階層			
		1千万円以上 5千万円未満	5千万円以上 1億円未満	1億円以上 10億円未満	10億円 以上
鉄鋼	売掛金	0,6891	0,8915	0,8241	0,1092
	流動資産	0,8252	1,1858	0,4758	0,0252
	固定資産	0,7260	1,1560	0,6721	0,0119
	資産合計	0,6361	0,9955	0,6053	0,0112
	買掛金	0,6440	0,9330	0,7977	0,0856
	流動負債	1,6284	0,9842	0,4317	0,0513
	固定負債	0,9380	1,0266	0,6033	0,0182
	自己資本	0,7959	0,9983	0,9029	0,0963
	売上高	0,8632	1,0358	0,7956	0,0333
	従業員	1,2200	1,0360	1,4225	0,0361

第4表 b)　主要費目に関する R^2 の変化（総資本）
(46年度法人統計年報)

産業＼資本金＼規模コード	1 200万円未満	2 200万円以上 500万円未満	3 500万円～ 1000万円	4 1000万円～ 5000万円	5 5000万円～ 1億円	6 1億円～ 10億円	7 10億円以上
農業	0,5762	1,5322	1,0545	4,2358	1,3990	0,6622	
林業	0,4158	0,7735	1,1536	1,2398	1,7345	1,0222	
漁業	0,9306	0,9201	0,8021	1,6808	0,8019	0,5292	0,0898
鉱業	2,0508	1,5364	1,0426	0,6317	1,2470	1,0916	0,2075
石炭鉱業			1,0000		1,1597	2,7242	1,6538
建設業	0,9023	1,0014	0,9498	0,5452	0,7969	0,3580	0,1198
食品製造業	1,2072	0,7832	0,9594	0,6567	0,9399	0,7330	0,4069
繊維工業	1,0769	0,7435	1,0423	0,4632	1,0594	1,6136	0,0757
紙パ製造	0,7673	0,9801	0,7670	1,0078	1,2538	0,8773	0,1299
化学工業	1,3604	0,7497	0,7656	0,6234	0,9700	1,3869	0,0641
窯・土石	1,3451	1,2412	1,0208	0,8812	0,9081	0,5333	0,1141
鉄鋼業	1,0928	1,2232	1,0640	0,8063	0,8354	0,9081	0,0095
非鉄製造	3,8862	1,2371	1,1950	0,7574	1,2050	1,0820	0,0562
金属製品	1,6783	0,8271	1,1420	0,5995	1,2004	0,6443	0,1650
一般機械	1,5528	0,8150	1,1148	0,7627	1,0023	1,0886	0,1787
電気機械	0,4915	0,8713	1,1293	0,4959	0,8273	0,7508	0,0130
輸送用機械	1,3340	0,9247	1,1293	0,6206	1,1747	0,5904	0,0457
精密機械	0,8022	0,9024	0,8389	1,0932	0,7191	0,8234	1,1387
武器製造	2,3378	1,2519	0,8349	0,4870	0,6850	1,2477	0,0968
その他製造	0,8548	1,0034	0,8963	0,7356	0,8982	0,9409	2,2293
卸売業	1,6161	0,9401	0,9413	0,6890	0,8970	0,5784	0,1114
その他の小売	0,9476	0,9185	0,9043	0,8519	1,0741	0,4655	0,9499
不動産業	0,8405	0,8427	0,9716	1,4277	1,2200	1,1438	0,3182
その他の不動産	1,6377	1,2562	1,2106	1,0107	1,0884	1,4276	0,1886
道路運送	1,0790	0,9954	0,9619	0,4989	0,7983	0,7178	0,1324
航空運輸	0,7244	0,7305	1,0799	0,5802	1,3195	0,7672	0,0433
その他の運輸	1,6221	0,6152	1,0868	0,9870	0,8159	0,3485	0,0271
電気業						2,0237	0,0063
ガス業				1,3119	1,2331	0,2443	0,0123
娯楽業	0,9225	0,8370	1,0567	0,9289	0,9562	0,6745	0,3255
放送業						0,9519	0,6865

第4表 b) 主要費目に関する R^2 の変化（固定資産）
（46年度法人統計年報）

規模コード / 資本金 / 産業	1	2	3	4	5	6	7
	200万円未満	200万円以上 500万円未満	500万円以上 1000万円未満	1000万円以上 5000万円未満	5000万円以上 1億円未満	1億円以上 10億円未満	10億円以上
農業	0,5970	1,5642	0,8721	0,5035	1,3352	0,5831	
林業	0,8329	0,7338	1,0708	1,1268	1,4398	0,9905	
漁業	0,9734	0,7888	0,8248	1,5909	0,7579	0,2980	0,0230
鉱業	2,2787	1,0880	1,0795	0,8487	1,2189	1,1129	0,3269
石炭鉱業			1,0000		1,1985	2,8056	1,6184
建設業	1,0719	1,0705	1,0073	0,5463	0,8507	0,4984	0,0580
食品製造業	1,6444	0,8998	0,9680	0,6877	0,9339	0,5746	0,1938
繊維工業	1,4023	0,7541	1,1583	0,6122	0,9587	1,6568	0,0986
紙パ製造	0,6387	1,1611	0,9951	0,9020	1,0368	0,8996	0,1590
化学工業	1,7710	1,0393	0,7202	0,6993	0,9971	1,4828	0,0782
窯・土石	1,3411	1,1741	1,1264	0,8231	0,9242	0,5079	0,0783
鉄鋼業	2,2856	1,1276	1,0083	0,8489	0,8664	0,9351	0,0105
非鉄製造	2,0679	1,4487	0,7400	0,7921	1,2104	0,6769	0,0610
金属製品	1,3462	0,9298	1,1796	0,7236	1,2515	0,6502	0,1702
一般機械	1,1440	0,7942	0,9706	0,8319	0,9115	0,6792	0,2156
電気機械	0,5685	0,8048	1,0407	0,5423	0,8701	0,7149	0,0276
輸送用機械	1,2962	0,7996	1,0407	0,7144	1,1968	0,6190	0,0635
精密機械	0,9733	0,7312	0,8230	1,1463	0,7236	0,6804	1,7520
武器製造	1,4305	1,2376	0,8734	0,5600	0,6705	1,2800	0,0244
その他製造	0,9050	1,1215	0,9562	0,6767	0,9283	0,9309	1,9330
卸売業	2,8328	0,8808	0,8662	0,6628	0,8434	0,6416	0,1003
その他の小売	0,7170	1,0167	0,9370	0,7528	0,9505	0,3863	1,1219
不動産業	0,9833	1,2805	0,7905	0,7393	1,2280	0,9260	0,2159
その他の不動産	1,6480	1,3141	1,2053	1,4123	1,0883	1,3907	0,2829
道路運送	1,0887	0,9258	1,0204	0,5604	0,7863	0,5924	0,1352
航空運輸	0,7377	0,7609	1,0917	0,5607	1,3173	0,5595	0,0444
その他の運輸	1,5417	0,6698	1,1066	0,9178	0,8661	0,4652	0,0262
電気業						1,6857	0,0078
ガス業				0,8871	1,2396	0,1835	0,0075
娯楽業	0,9395	0,9020	0,9764	1,0020	0,8707	0,7919	0,1063
放送業						1,0025	0,4053

第4表 b） 主要費目に関する R^2 の変化（売上高）
（46度法人統計年報）

規模コード / 資本金 / 産業	1	2	3	4	5	6	7
	200万円未満	200万円以上 500万円未満	500万円～ 1000万円	1000万円～ 5000万円	5000万円～ 1億円	1億円～ 10億円	10億円以上
農業	0,5245	1,3292	1,1046	0,9698	1,3588	0,9769	
林業	0,4122	0,7889	1,1222	1,4525	1,6232	1,0365	
漁業	1,6237	1,0329	1,2043	1,2537	0,8644	0,3897	0,0055
鉱業	2,4593	1,5273	0,9923	0,7752	1,2753	1,3102	0,2382
石炭鉱業			1,0000		1,1939	1,5408	1,5308
建設業	1,0513	0,9393	1,0145	0,5065	0,8217	0,5001	0,1744
食品製造業	1,0612	1,0612	0,9968	0,9280	0,9377	0,7638	0,2891
繊維工業	0,7931	0,8122	1,0842	0,4980	0,9914	1,5153	0,1037
紙パ製造	0,9260	1,0088	0,7495	1,0641	1,2961	0,9585	0,2438
化学工業	2,0869	0,9992	0,8110	0,6281	1,0104	1,3301	0,1316
窯・土石	1,0478	1,2424	1,0617	1,0744	0,8285	1,0722	0,2788
鉄鋼業	0,8985	1,2017	1,1554	0,7600	0,9127	0,9376	0,0252
非鉄製造	5,7385	1,0264	1,4616	0,7264	0,7996	1,6429	0,2213
金属製品	1,4957	0,7551	1,0285	0,6197	1,2367	0,7674	0,2534
一般機械	1,5486	0,8366	1,1448	0,7963	0,9260	1,2621	0,2977
電気機械	0,8533	0,8705	0,9757	0;5588	0,8814	1,0900	0,0627
輸送用機械	1,1940	0,9735	0,9757	0,7520	1,1693	0,8032	0,0799
精密機械	0,8309	0,8997	0,9283	1,2722	0,7459	1,0293	2,1571
武器製造	2,0531	1,2016	0,7751	0,5212	0,5871	0,9426	0,0603
その他製造	1,0421	1,0794	0,9402	0,7503	0,8859	1,0103	2,1592
卸売業	1,4468	0,9526	0,9374	0,8113	0,8331	0,6891	0,1103
その他の小売	0,9305	0,9893	0,9500	1,0601	1,0522	0,5227	0,8467
不動産業	0,8697	0,9741	1,1718	1,2455	1,1880	1,1771	0,4946
その他の不動産	0,8128	1,1018	1,0269	0,8051	1,0535	1,3770	0,7016
道路運送	1,0063	0,8938	1,0454	0,4616	1,0034	0,7369	0,1217
航空運輸	0,5813	0,7309	1,2326	0,9476	1,2716	1,3266	0,1807
その他の運輸	1,9035	0,6172	1,0351	1,3723	0,9122	0,6198	0,0557
電気業						1,2590	0,0105
ガス業				1,8232	1,2346	0,4434	0,0138
娯楽業	1,0224	1,0090	1,0065	0,8592	0,9254	0,5439	0,1640
放送業						0.6068	0,9408

第4表 c) $R^2=\frac{\sigma_x^{2\prime}}{\sigma_x^2}$ の全項目平均の変化

(46年度法人統計年報)

規模コード	1	2	3	4	5	6	7
産業 \ 資本金	200万円未満	200万円以上 500万円未満	500万円～ 1000万円	1000万円～ 5000万円	5000万円～ 1 億 円	1 億 円 ～ 10 億 円	10億円以上
農業	0,5864	1,1008	1,0675	1,0726	1,3407	0,8701	
林業	0,7240	0,7985	1,0800	1,2630	1,4872	1,0230	
漁業	1,0952	0,9908	0,9377	1,3780	0,8701	0,6565	0,1327
鉱業	2,0676	1,3706	1,0005	0,7671	1,2047	1,1463	0,3588
石炭鉱業			1,0003		1,2574	2,3084	1,5097
建設業	0,9978	0,9888	0,9795	0,6157	0,8446	0,4526	0,2812
食品製造業	1,1395	0,8525	0,9610	0,8291	0,9631	0,7550	0,5263
繊維工業	1,0178	0,8179	1,0328	0,5810	1,0347	1,5468	0,1243
紙パ製造	0,9455	0,9673	0,8134	1,0235	1,2212	0,9265	0,2165
化学工業	1,5632	0,9397	0,8338	0,6642	0,9720	1,3709	0,1262
窯・土石	1,3934	1,1759	1,0227	0,9407	0,9166	0,8100	0,3355
鉄鋼業	1,3214	1,1361	1,0469	0,8589	0,8679	0,9520	0,0345
非鉄製造	3,4339	1,1534	1,2040	0,8253	1,1343	1,1373	0,1858
金属製品	1,6581	0,8253	1,1053	0,6748	1,1686	0,7669	0,2606
一般機械	1,4563	0,8527	1,0989	0,8278	1,0025	1,0818	0,2479
電気機械	0,5986	0,9102	1,0191	0,6068	0,8710	0,9098	0,0392
輸送用機械	1,2095	0,9480	1,1159	0,7005	1,0995	0,7652	0,0730
精密機械	0,8839	0,9100	0,8636	1,1181	0,7592	0,9298	1,1643
武器製造	1,5879	1,0477	0,9483	0,5760	0,7119	1,1563	0,1226
その他製造	0,9651	1,0631	0,9492	0,7967	0,9947	1,0126	1,8768
卸売業	1,7510	0,9241	0,9382	0,7136	0,9109	0,7037	0,1894
その他の小売	0,9643	0,9702	0,9287	0,9811	1,0669	0,6923	1,1079
不動産業	0,8858	0,9101	0,9642	1,1460	1,0798	1,1530	0,5157
その他の不動産	1,3249	1,1150	1,1415	0,8975	1,1040	1,2755	0,3903
道路運送	1,1412	1,0019	1,0058	0,5799	0,8833	0,8428	0,1829
航空運輸	0,7324	0,7718	1,1431	0,7185	1,2466	1,0413	0,1352
その他の運輸	1,5001	0,7017	1,0559	1,1775	0,8828	0,4608	0,1361
電気業						1,3606	0,0367
ガス業				1,3276	1,1966	0,4225	0,1288
娯楽業	0,9469	0,9287	1,0555	0,9944	0,9965	0,6883	0,6205
放送業						0,8904	0,8432

しないのである．以上で明かなように，経験資料に於ても前節の結論 10 が略妥当するのである．すなわち相関係数の標本抽出法の選択に対する役割は，状況により可成りまちまちであり，限定された基準を与えるに過ぎない．それは規模階層に依存する．事実これまで扱って来た歪みのある理論的諸分布及び経験的諸分布に対して共通していることは，大規模層に対しては，比較的規模比例確率抽出法が有利であるという見方であろう．他方より有効な基準を与えうるものとして

問題点 1. 『(57) 式の抽出確率の決定に於ては対数相関係数．対数回帰係数或いはそれに準ずる諸係数が有効であることが期待される．従って今後其等を既存資料或いはより一般的な見地によって再検討することが今後の一つの課題といえるであろう．且つその時もし其等の有効性が立証されるならば，従来の間分散最大化原理にかえて，対数相関係数の極大化，或いは対数残差極小化を原理として層化し，又同一原理によって最適割当法を発展的に適用することが可能となろう．』

8. むすび法人統（計の為の標本設計の提案）

以上各節における理論的実際的諸結果に照して，現状に則した標本設計として，次の混合形態が適当と思われる．即ち

提案，法人を産業別資本金階層別に層化した後，以下の抽出形態に従って最適割当法を適用する．

資本金階層	万円 万円 200～499	万円 万円 500～999	万円 万円 1000～4999	万円 万円 5000～9999	億円 万円 1～99999	億円 10～
抽出形態	等確率抽出	〃	〃	〃	資本金規模比例確率抽出	悉皆抽出

処で抽出作業の実際に於ては，資本金の累積和に対して systemmatic sampling を施行することになるから，或程度の偏りが生ずることは免れない．第 6 表及び第 7 表は，資本金 1 億円以上 10 億円未満法人層に対する其の意味を含めた実験例であるが，全体を通じて一致性に関しても規模比例確率抽出法が，等確率抽出法より好ましい結果を与えているといえるであろう．

× × ×

本稿の掲載した諸資料は，大蔵省証券局資本市場課の方々の御協力によるものである．
又，本稿の計算，製図，製表等については河合啓子嬢，太田芳子嬢の助力によるものである．
比処に併せて謝意を表したい．

参 考 文 献

[1] Aitchison, J. and Brown, J.A.C. '*The Lognormal Distribution*,' Cambridge at the university Press (1957).

[2] Hansen, M.H. and Hurwitz, W.N. 'A new sample of the population: sampling principles introduced into the Bureau's monthly reports on the labor force,' U.S. Bureau of Census (1944).

[3] Horvitz, D.G. and Thompson, D.J., 'A generalization of sampling without replacement from finite universe,' Journal of the American Statistical Association, Vol. 47 (1952) pp 663–685

[4] Mardia, K.V. 'Multivariate Pareto distribution,' Ann. Math. Statist. Vol. 33 (1962), pp 1008–1015

[5] Raj, D. 'The use of systemmatic sampling with probability proportionate to size in a large scale survey,' American Statistical Association Journal, march. (1964), pp 251–255

第5表　資本金額1億円以上10億円未満層に関する抽出実験結果表（その1）
（49年度法人企業統計季報1～3月分）

業種	項目	全数 a	規模比例推計 b	単純推計 c	実現率(%) b／a	実現率(%) c／a	母変動係数	累計社数 a	累計社数 b	累計社数 c
鉄鋼	資本金	42,387	42,387	41,682	100.0	98.3		145	88	72
	売掛金	240,577	242,351	245,238	100.7	101.9	0.939			
	棚卸資産	109,600	107,304	111,960	97.9	102.2	0.857			
	固定資産	314,930	326,749	338,562	103.8	107.5	0.943			
	資産合計	849,041	863,092	895,942	101.7	105.5	0.776			
	買掛金	328,257	339,566	327,348	103.4	99.7	0.927			
	短期借入金	65,362	65,336	77,008	100.0	117.8	1.489			
	長期借入金	166,907	158,834	186,340	95.2	111.6	1.509			
	営業損益	34,362	35,426	33,636	103.1	97.9	1.190			
	売上高	300,357	303,974	294,764	101.2	97.9	0.793			
（145社）	従業員数	51,319	46,935	56,662	91.5	110.4	1.292			
非鉄	資本金	32,622	32,622	32,988	100.0	101.1		102	67	51
	売掛金	174,337	149,718	210,992	85.9	121.0	1.534			
	棚卸資産	100,381	91,006	104,394	90.7	104.0	1.121			
	固定資産	138,643	132,482	150,594	95.6	108.6	0.622			
	資産合計	513,828	465,476	591,670	90.6	115.1	1.032			
	買掛金	207,721	189,642	250,220	91.3	120.5	1.261			
	短期借入金	45,719	33,908	61,666	74.2	134.9	2.063			
	長期借入金	73,490	75,598	90,942	102.9	123.7	0.938			
	営業損益	25,808	22,485	30,458	87.1	118.0	1.219			
	売上高	227,530	192,722	269,426	84.7	118.4	1.726			
（102社）	従業員数	33,218	32,615	39,084	98.2	117.7	0.851			
金属製品	資本金	64,644	64,644	65,952	100.0	102.0		348	141	124
	売掛金	325,150	302,616	302,964	93.1	93.2	0.914			
	棚卸資産	246,310	218,830	236,444	88.8	96.0	1.133			
	固定資産	366,777	355,729	344,922	97.0	94.0	0.801			
	資産合計	1,199,693	1,129,818	1,135,494	94.2	94.6	0.770			
	買掛金	455,315	412,115	420,300	90.5	92.3	0.886			
	短期借入金	124,784	115,426	116,254	92.5	93.2	1.462			
	長期借入金	201,198	195,856	184,762	97.3	91.8	1.172			
	営業損益	45,664	44,626	40,026	97.7	87.7	1.075			
	売上高	396,599	363,394	365,318	91.6	92.1	0.892			
（248社）	従業員数	104,686	98,093	99,562	93.7	95.1	0.084			

第5表 (その2)

業種	項目	全数	規模比例推計	単純推計	実現率(%)		母変動係数	累計社数		
		a	b	c	b/a	c/a		a	b	c
	資本金	89,187	89,187	89,522	100.0	100.4		327	189	164
	売掛金	545,012	580,365	626,646	106.5	115.0	1.370			
機	棚卸資産	366,755	349,806	407,328	95.4	110.1	1.048			
	固定資産	348,403	335,948	346,626	96.4	99.5	0.804			
	資産合計	1,613,828	1,623,056	1,739,844	100.6	107.8	0.979			
	買掛金	566,153	566,093	638,376	100.0	112.8	1.346			
	短期借入金	264,879	272,861	292,026	103.0	110.2	1.418			
械	長期借入金	217,864	203,401	230,386	93.4	105.7	1.437			
	営業損益	40,608	43,236	42,996	106.5	105.9	1.483			
	売上高	462,122	462,855	505,478	100.2	109.4	1.324			
(327社)	従業員数	139,137	131,164	144,392	94.3	103.8	0.846			
	資本金	77,144	77,144	79,196	100.0	102.7		289	171	144
	売掛金	319,744	355,866	305,296	111.3	95.5	1.087			
電	棚卸資産	292,178	301,228	295,006	103.1	101.0	0.797			
	固定資産	301,376	320,529	300,712	106.4	99.8	0.778			
気	資産合計	1,168,403	1,249,889	1,146,156	107.0	98.1	0.743			
	買掛金	425,181	479,497	394,516	112.8	92.8	1.223			
機	短期借入金	163,513	164,082	168,518	100.3	103.1	1.195			
	長期借入金	139,785	146,341	142,058	104.7	101.6	1.262			
器	営業損益	30,572	30,670	26,554	100.3	86.9	1.597			
	売上高	418,197	464,424	396,062	111.1	94.7	0.979			
(289社)	従業員数	173,585	194,552	172,572	112.1	99.4	0.983			
	資本金	67,973	67,973	69,154	100.0	101.7		231	141	116
輸	売掛金	387,641	342,586	377,596	88.4	97.4	1.147			
	棚卸資産	279,570	239,068	297,940	85.5	106.6	1.807			
送	固定資産	493,889	453,559	528,550	91.8	107.0	1.262			
	資産合計	1,550,520	1,313,298	1,643,984	84.7	106.0	1.770			
用	買掛金	507,724	466,706	513,368	91.9	101.1	1.082			
	短期借入金	165,152	147,703	188,750	89.4	114.3	2.127			
機	長期借入金	277,786	228,285	292,668	82.2	105.4	2.511			
	営業損益	27,652	24,499	28,536	88.6	103.2	1.377			
械	売上高	500,612	461,118	542,350	92.1	108.3	0.966			
(231社)	従業員数	161,766	154,387	170,044	95.4	105.1	0.782			

第6表　資本金額1億円以上10億円未満層に関する推計方法の差による産業平均実現率の比較

	規模比例推計		$\frac{1}{2}$ 抽出単純推計		
	平均実現率(%)	実現率変動係数	平均実現率(%)	実現率変動係数	
売掛金	99.6	0.093	104.4	0.108	① 49年1～3月期，農林水産，電力，ガスを除く資本金1億円以上10億円未満のデータ
棚卸資産	100.9	0.099	104.5	0.107	
固定資産	101.5	0.062	101.7	0.119	
資産合計	100.0	0.070	103.1	0.107	
買掛金	99.3	0.086	99.4	0.106	
短期借入金	102.8	0.137	107.5	0.150	
長期借入金	97.9	0.064	104.4	0.179	② 規模比例；抽出間隔4億円
営業損益	100.3	0.121			
売上高	100.2	0.078	101.5	0.105	
従業員数	102.0	0.109	103.8	0.113	③ 実現率＝推計値／全数値
資本金	100.0	0.000	100.2	0.056	

[6] Rao, J.N.K., 'On the estimate of the variance in unequal probability sampling, Ann. Inst. Statist. Math., Vol. XIII (1961), No. 1 pp 57–60
[7] 田口時夫： 所得分布の多次元的性質（予定）

備　　考

1. 昭和49年度科学研究費補助金とは，昭和49年度試験研究，課題番号983001のことである．
2. 46年下期，47年上，下期，48年上期の法人統計資料の利用に関しては
　行管承統第399号，蔵証第2894号
によるものである．
又48年度法人統計資料の利用に関しては，
　行管承統第267号，蔵証第1446号
によるものである．
3. 紙面の都合により，資料は全面的に掲載する事が出来なかった．本稿に掲載された資料と同種の内容であって且割愛された資料についても，全体として同様な傾向が認められたことを附記する．

正　誤　表

統計数理研究所彙報 第22巻 第1号 1974
p 18　下段より10行目及び11行目

	正	誤
10 行目	勾配の変化の少ない点	勾配の変化の著しい点
11 行目	微係数が最小となる	微係数が最大となる