統計数理 第34巻 第2号（1986）

標本調査論の再検討*

―― 有意選出法の基礎について ――

統計数理研究所　田　口　時　夫

（1986年3月　受付）

目　　次

はしがき

標本調査論は戦後我国で戦後事情に伴い著しく普及し，昭和20年代に於いてほぼその実務及び理論の体系が完成されるに至った．

其後の長期に亙る高度成長期に於いては，この整備された統計体系の下で，統計調査は量的に増大したが，特に質的向上は認められないのである．然るに石油ショック以降，特に昭和50年代に入り，構造変化を生ずるとそれは統計環境にも及び，任意抽出論における停滞も加わって，標本調査論をその根底から再検討する現実的な契機を形成している．

振り返ってみると戦後我国の標本調査体系は専ら確率論に基礎をおくフィッシャーの標本論以来の任意抽出法の流れを汲むものであった．近代化に一歩遅れをとった我国の歴史事情からして，当時支配的であった任意抽出法の摂取・吸収に追われたのは，やむを得ないことである

* 本稿は統計数理研究所の60年度共同研究計画の一環として行われた「官庁統計の調査・集計・解析に関する研究」(60-共研-23 代表者田口時夫)の成果の一部を示すものであり，昭和61年3月17日における同名の研究集会に於ける報告及び61年度の第54回日本統計学会に於いて「有意選出法について」と題した報告を主な内容としている．

キーワード：代表法，回帰関数の弾力性係数(弾性値)，定――，変動――，不等確率抽出法，統一理論，有意選出法，典型調査，集中解析，回帰適合曲線，マルディアの多変量パレート分布，ノン・ストカスチック な推論，集団構造，集中構造，相互依存関係

が，今日の統計環境では，最早それのみに終止することは許されないことといえよう．つまり，独自の開発を進めることが要請されるのである．その為には我々は標本調査の歴史を遡り，その形成の論理を辿り，将来への展望を図らねばならない．[註1)]

ところで問題は単に我国のみの事情にとどまらない．ハンセンによる不等確率抽出法の提案以降，標本抽出の統一理論が発生し，そこでは有意抽出法の再検討を促すような結果も少からず見受けられるのである．思うに任意抽出法と有意選出法をめぐる代表法の論争は，世上一般ではジーニに対するネイマンの批判によって結着をみたかの如く理解されている．然し乍ら彼等の論文を再検討し，一方でデータ解析を進めると，そこに若干の疑点が生じて来る．筆者は本稿に於いて，一方では統一理論の成果を批判的に摂取し乍らも，原則的にジーニの方向に沿ってそれを補完し，発展させる方向で有意選出法を再提起する．それと共にそれを基礎づける方向として回帰適合曲線を提案し，此等の方法を規定する基本的性格が各種の弾性値によって決定される事を示すであろう．

註1) 戦後の我国の統計におけるSturm und Drangを松下嘉米男はこう述懐している．「終戦後，アメリカの文献を見る事が出来るようになったが，それによって戦時中アメリカに於いては，統計理論がその応用と共に非常に発達していたことが判った，当時，研究所に於いてもこれらアメリカの成果が紹介されることがしばしばあった．その興味をもった主なものはWaldがその理論を確立したSequencial analysis並びに，これにまつわる事柄等，Neyman, Hansen, Hurwitz, Madowによって発展したSampling理論並びにその応用等であった.」「一方，応用方面に於いてもSampling理論を応用することが研究所の指導の下に，各方面で行われる様になった．これは当時，G.H.Q.の影響も多分にあって，諸機関がSampling理論の方法を取り入れる様になったためである．今日我が国に於いて，新しい意味の統計理論が非常に重要視される様になったことには，G.H.Q.並びにその後日本に来たアメリカ統計学者の影響がすこぶる大きかった事を認めねばならないであろう.」以上松下嘉米男『十周年に当りて』統計数理研究所彙報，第2巻第1号1954年6月p.3による．

序　　論

標本調査法は，任意抽出法と有意選出法を含む代表法として，1895年に国際統計協会ベルン会議に於いて，キェールにより提起された．少くとも国際的な会議の議題として提起されたのは，この時点であるという．(馬場 (1969), p.192) それは「従来，統計学界で主流をなしていたのは，全数調査を万能とし，個別典型調査を補助とみたものであった」ことに対処して行われたものであった．「キェールは，このように国際統計会議において，重要視してきた典型調査法に対し，新たに代表調査法を提唱したのである.」(馬場，上掲書p.194) それについてキェールは1899年国際統計協会ブタペスト会議の報告において「ある地区とか，一都市のある街区の，詳細な研究は代表調査 (une investigation representative) ではない．もしいま問題としている地区とか，街区とかが，典型(type)と考えられるならば，それは典型調査(une investigation typologique) である．… 代表法は散らばっている多数の観察単位 (地区) を必要とする．そして全集団のなかで種々の特性をもつ地区が，出来るだけ同じ割合で全集団(l'ensemble)を代表するように取らねばならない.」(同上pp.194-195) と説明している．

又ジェンセンは，1925年国際統計協会ローマ会議に於いて，統計調査法についての簡潔な説明を与えているが，それは極めて示唆に富むものである．それによると，

「(1) 部分調査(partial investigation)．これは知識を得たいと望んでいる全領域を，調査するのでなく，その領域の一部分のみを調査するとき，統計的調査のすべてを総称する．その調査部分は偶然的に入手されても，調査のために選定されても，いずれでもよい．(2) 代表法(representative method)．これは部分調査から得られた結果を一般化できるように，代表標本

を確保する方法で，次の二つにわかちうる．(3) 無作為選出法(random selection)．これは調査の主題や目的と結びつかない，何らかの機械的な原理にしたがって，全体における各単位の，標本に含まれる確率が等しいように，選出の仕方を考えて，多数の単位を選ぶ方法．(4) 有意選出法(purposive selection)．これは多数の単位群を選出するとき，選出された群が相集って，すでに統計的知識となっている諸特性(characteristics)について，全体とできるかぎり近い平均•比率を生ずるように行う選出方法．(5) 標本(sample)．これは無作為選出法にしろ，有意選出法にしろ，調査のために選ばれた，単位の群の代表集団である．

ジェンセンは社会的集団現象の観察につき，先ず簡略なものとして，問合せ法(inquiry method, アンケート)をあげている．これは調査事項に委しい多くの人々に，問合せをして，彼等の個人的な経験や観察に基づいた，主観的な見解を求め，これをもとにして，調査者が一つの総合的な報告を作る場合である．次に彼は，記録法(monographic method)をあげる．これは集団の構成部分のうち，典型的(typical)と思われる部分について，客観的な調査をなす場合である．すなわち典型調査はこれである．」(前掲書 pp. 148-149, A. Jensen (1926))

我々は本稿でその詳細を述べる訳にはいかないが，有意選出法と任意抽出法の並存を認める代表法の観点に立ち戻ることは極めて重要であると考える．[註1] この観点の下で我々は概ね第1表のような，一つの系譜を与えることが出来る．つまり此の表に収めた各国の理論的•実践的指導者たちは，それぞれの方法について重要な規定を具体的な形で与えているからであり，それによって我々は二つの方法の並存を認める代表法への回帰の必然性を読み取ることが出来るからである．残念乍ら本稿でその詳細を与えることは出来ないので，ここではそれぞれの代表者の論文を通して，その概要を第2表の形式で示すことにする．

勿論こうした縮約は史家により多くの指弾をうけるであろうが，本稿第1章及び第2章の具体的提案を主眼とした本稿では，両方法が相互に補完的な立場にあり，絶対的な対立や，一方への従属といった関係にはないことを示し，将来への大掴みな展望を与えることが出来れば，充分であると考える．従ってキェールとジェンセンを第I期に位置づけ，第II期にはジーニ，ガルヴァーニとアンダーソンを含めることにした．又任意抽出法に関しては，フィッシャー，ボーレー，ネイマンを何れも第I期に，ハンセンとフルヴィッツを第II期として捉えた．第III期は，其後の今後に至る諸々の研究者の成果を概括する意図で構想したものであるが，有意選出法の此期に該当する方法は他に未見であるので，敢えて予想としたのである．但し本稿第1章及び第2章で示す方法は，相互に異なる系譜に従うが，共に此期に属するものとして意識的に第III期を構成している．

重要なことは，或期の有意選出法の構図が期をずらして任意抽出法の基盤の下で摂取され，或場合はストカスチックな手法が次期の有意選出法を生む引金となっていることである．例えばジーニ，ガルヴァーニはストカスチックな最小二乗回帰直線を多分ボーレーから導入して自己の方法を表現したのであり，逆にハンセン，フルヴィッツは有意選出法にならってコントロールを導入して任意抽出法の改良を果たしたのである．つまり此の相互作用があってこそ，それぞれの方法に進歩と発達がみられるのであり，常に一方を主とし他方を従とする傾向として標本論史を捉えることは我々のとらないところである．そのような傾向は存在しても，部分的なものであり或いは一時的なものだからである．又そのような一方通行的な理解は寧ろ自己否定に導き，永続性をもたないように思われる．

然し両者の並存と相互作用を認めるといっても，両者の間には基本的な対立が存在するのである．それは第3表のように，それぞれが指摘されて来た問題点に於いて捉えることが出来よう．

ところで，第1, 2, 3の各表のように水平化した表現は一つの概観を与えるものではあるが，

第1表 代表法の系譜

有意選出法	任意抽出法
キェール（ノルウェー） 国際統計協会ベルン会議（1895）	フィッシャー （1920年代発足）
ジェンセン（デンマーク） 国際統計協会ローマ会議（1925）	ボーレー（イギリス） 国際統計協会ローマ会議（1925）
ジーニ，ガルヴァーニ（イタリア） カイロ国際統計機関会議（1927）	ネイマン（ポーランド，アメリカ） 王立統計協会（1934）
アンダーソン（ブルガリア） 王立統計協会（1934）	ハンセン，ハーヴィッツ（アメリカ） アメリカ・センサス局（1944）

第2表 キーワード

	有意選出法	任意抽出法
第I期	調査区選出，区内全数，順位付け，系統選出，コントロール	個体，等確率抽出，層化法
第II期	直線回帰，コントロールの母平均と標本平均の一致性	変動確率抽出
第III期	曲線回帰，弾性値（予想）	モデルの導入，最適化

第3表 問題点

有意選出法	任意抽出法
1. 主観性が強く，一般的に偏りも伴う．	1. 繰り返し操作可能を前提としている．
2. 全般的に数学的根拠が薄い．	2. 信頼性に関して諸説がある．
3. 精度表現がない．	3. 層化基準は必ずしも客観的でない．
	4. 乱数という人為的且つ不統一で実現困難な手続き及び概念を必要とする．

方法の発生や交替の内的動因を伝えるには不充分である．我々はこうした方法概念の形成・変化について特に画目すべき論文に触れねばならない．

その第1は，有意選出法の第II期に位置づけたジーニ，ガルヴァーニの論文である．彼等が標本抽出を試みたのは，地域間の比較や，その内部分析等の精密作業を行うのを可能にする為に，人口，社会，経済，地理的諸特性について全国を代表するような「一般的標本」を得る事であった．具体的に，

「出生率の順序で整理したリストから順番の番号が1の数字で終わる郡 —— すなわち，1,11, 21,…番を占めた郡 —— を選抜した．ただし，標本にはイタリアのすべての州とともに，富・産業上の特徴・海抜の故にイタリアの大・中都市全体を充分に代表すると思われる大・中都市も含まれることが望ましかったので，最初に選んだ郡の他にいくつかの郡を加えた方が良いと判断した．かくして，27郡から形成された暫定的標本を作成することに達した．この標本は出生率の点では最高に代表的なものであった．

ところが，予想しうるように，広さの点では，予め規定した量的基準と一致したこの標本は，選抜の際に考慮しなかった諸特性に関しては，当初の一見で母集団の一般的表象のために要求されうる最低の条件を満たすことから相当遠いものになった．」(p. 8)

その為に彼等は出生率以外の特性を加味し，経験に基づいた修正を加えるといった基本的に

単純で理論的に不完全な作業を行ったのである．然しその過程及びその後の推計に当って，彼等はそれまで主観性が強く数学的根拠に乏しいといわれた有意選出法に不充分ではあるが統計的解析を加え，数学的方法の適用に努力し，その可能性を探索したのである．[註2)]

彼等はその結果について最終的に必ずしも満足していないが，それは「一般的標本」という概念に含まれているように，あまりにも多目的，多面的な代表性を求めた結果でもある．一つの標本に対して全面的な代表性を求めることは，今日の如何なる調査法によっても困難であろう．

扨，第2の論文は，任意抽出法の第I期に位置づけたネイマンによるジーニ及びガルヴァーニの論文に対する批判的報告である．[註3)] 彼はここで強く，有意選出法は任意抽出法の一部であり，その特殊な場合であると結論づけている．然し乍ら此の主張は，有意選出法を客観的に把えた結果としてではなく，自己の主観的判断を加え，又彼の確率論的視野に収まるように変形を加えた結果であることは，次の二つの章句によって端的に伺うことが出来る．

つまりジーニとガルヴァーニは，彼等が有意選出法を提案するのは，その前に任意抽出法を試みて成功しなかったからであるがそれを具体的数字を用いて説明した箇所を引用した後，ネイマンは次のような註釈を施すのである．「しかしながら，いま見たように抽出単位を決めた経緯からすると，抽出の原則は予め決められていたように思われる．」(p. 99 註3参照)従ってここでは当初からの疑惑と不信の感情を読みとることが出来る．更に内容に立ち入ると，ジーニとガルヴァーニの用いたコントロール（出生率）により順位づけた後の系統選出を，コントロールによる層化であり，層内から調査区の一部を無作為に抽出することであると，彼流の解釈を与えた際に再び註釈が加えられる．つまり次のような弁明を試みるのである．「確率論的に有意選出法を扱おうとすると，このように解釈しなければならないことは強調しておく．ランダムな変動或いは無作為抽出でないときは，確率誤差等の確率の議論は一切出来ない．」(p. 109 註3参照)つまり彼においては初めから確率論が鉄則なのであり，それ以外のことは考えられないのである．我々はネイマンの確率論に対する傾倒の深さには打たれるが，有意選出法の統計目的に対する客観的科学的判断を求める立場からは，彼の論文は何も寄与するところがないという他はない．

事実，この報告によって一応の成功を得た（ネイマンは自ら「有意選出法に対する私の批判は充分説得力があった」(p. 139 註3参照) と満足げに答えている．）ネイマンは，これを一つの契機として其後アメリカで活躍するのであるが，この報告後僅か10年にして同じセンサス局のハンセン及びフルヴィッツによって次のように訂正されるのである．「しかし，単純任意抽出法を使わないで有意選出法の生む偏りを避けると同時に，単純任意抽出法よりも大幅に標本誤差を減少させるような特別な原理が導入された．それは任意抽出法を行うのだが，1次抽出単位が選出される確率はその規模既ち1940年代における人口に比例させるというものだった．」[註4)]

ここでネイマンが任意抽出法の原則とした等確率抽出は斥けられ，他方有意選出法を特徴づけるコントロールが層化の為ではなく，抽出確率そのものに導入されるのである．しかも彼等の提案は有意選出法を意識して，その長所を取り入れたとみて間違いないであろう．ネイマンの批判にも拘らず，彼等には任意抽出法と有意選出法が依然として対置され，並存されていたとする他はない．ハンセン及びフルヴィッツの提案は同時に両者を統合もしくは接近させる試みでもあった．それは我々に，彼等の論文を画目すべき第3の論文とするに足る充分な根拠を与えるものであろう．

註1） キェールやジェンセンの有意選出法の紹介は馬場（1969）に詳しい．

註2） Gini and Galvani (1929) 参照．此の論文の検討は一橋大学大学院の Carrado Molteni 氏の協力に負うものである．Gini と Galvani は又註として次のように言っている．「抽出の単位として市町村を採

用できなかった理由を別にして，任意抽出の方法の実際に限って言えば，8354の市町村から適切な数を抽出すれば，郡別の抽出より間違いなく良い結果が得られたであろう．さらに層別抽出方法——つまり全市町村を幾つかの「類似した」集団に分け，この各々の集団から同じ割合でいくつかの市町村を任意に抽出する方法——を以って標本を抽出したならば，より良い結果が得られたかもしれないが，多数の特性に対して同時に層別を行う時に如何なる困難にぶつかるかは誰にでも分かるであろう．なお層別抽出法に関しては，前掲の Bowley と Jensen の論文を参照」(p. 6)

註 3) Neyman (1934) 参照．この論文の検討は岸野洋久研究員（統計数理研究所）の協力によって行われた．

註 4) Hansen and Hurwitz (1944) p. 3 参照．此の論文の検討には岸野洋久研究員の協力を得た．

1. 変動確率抽出法におけるモデルの導入*

ハンセン，フルヴィッツ以降の任意抽出法は「地域抽出法」とか「人口規模に比例した確率抽出」といった具象性を脱却して，変動確率抽出法として抽象され，その一般理論の形成に向った．序論の第2表に掲げたように第III期の1つの特徴は，分散最小というネイマン基準の適用範囲の拡大にあると考えられる．それは単に標本の各層への割当や，推定方式の面に限らず，層の形成にもコントロールを導入してその最適分点を決定するといった試みや，変動確率抽出を最適化する試みによってその適用領域を拡大した．[註1)]

つまり此期に於いては，かつては有意選出法を特色づけるものであったコントロールが，任意抽出法に於いて当然の前提として導入された上で最適化が論じられている．そしてその結果の記述に於いて登場するのは一般に曲線相関・曲線回帰であり，弾性値であって，母集団の非線型構造に正面から取り組むことを要請するのである．このことを具体的に示す一例として変動確率抽出法に最適原理を導入した結果を示すことにしよう．

一般に不等確率抽出法は，補助変量 X が与えられた場合，その正値一価関数 $\varphi(X)$ に比例した抽出確率による標本抽出法と考えることが出来るが，此の抽出法によって得られた標本 $(X^{(i)}, Y^{(i)})$, $i=1, 2, \cdots, n$ を用いて目的変量 Y の母平均 $\mu_y=\frac{1}{N}\sum_{i=1}^{N} y_i$ を推定する為に，不偏推定量

$$\tilde{Y}(\varphi)=\frac{1}{n}\mu_\varphi \sum_{i=1}^{n}\frac{Y^{(i)}}{\varphi(X^{(i)})}\,;\ \mu_\varphi=\frac{1}{N}\sum_{i=1}^{N}\varphi(x_i) \tag{1.1}$$

を用いることにすると，その分散 $\sigma^2_{\tilde{Y}(\varphi)}$ はよく知られているように

$$\sigma^2_{\tilde{Y}(\varphi)}=\frac{1}{n}\left[\mu_\varphi\cdot\frac{1}{N}\sum_{i=1}^{N}\frac{y_i^2}{\varphi(x_i)}-\mu_y^2\right] \tag{1.2}$$

となる．ここで (1.1) の最適条件を与えるものとして，

定義 1.1 「φ の関数族に於いて $\sigma^2_{\tilde{Y}(\varphi)}$ を最少ならしめる $\varphi^{(0)}$ が存在するとき $\varphi^{(0)}$ に比例する抽出確率を Y の平均を推定する為に $\tilde{Y}(\varphi)$ を用いた場合の最有効抽出確率という．」

意味の最適抽出確率を用いることが出来るが，この場合次の定理が成立する．すなわち

定理 1.1 「目的変量 Y が正値で且つ補助変量 X に対して

---

* 本章は昭和60年度日本統計学会における共通テーマ「戦後の統計調査の問題点と将来のあり方〈国富調査〉」における報告を主体としている．

(1.3)
$$\log Y=\varphi(X)+\log\zeta\ ;\ \zeta \text{ は } X \text{ と独立}$$

となる構造をもつならば，$e^{\varphi(X)}$ 比例確率抽出が最有効抽出法を与える．又その場合 $\bar{Y}(\varphi^{(0)})$ の変動係数 $C^{(0)}$ は

(1.4)
$$C^{(0)}=\frac{1}{\sqrt{n}}\frac{\sigma_\zeta}{E(\zeta)}=\frac{1}{\sqrt{n}}C_\zeta$$

である．ここで C_ζ は ζ の変動係数を表わす.」[註2)]

構造に関する (1.3) の条件は Royall (1970a) のモデルに近い性質を示すが，この定理は二つの階層内の目的変量の総額がほぼ等しい時，標本数 n もほぼ等しくなり，従ってこの二階層のパーセント精度もほぼ等しいことを示すものである．これは階層別統計表の利用価値を高めるものであり，従来の層化任意抽出法によって得られない利点である．然もそれは非線形な回帰構造 (1.3) を前提として初めて得られるのである．

扨，数多くの経済統計の分析によって極めて一般的に認められている経験的な回帰関数は対数線形回帰形式

(1.5)
$$\log Y=\alpha+\beta\log X+\log\zeta$$

である．ここで β はしばしば Y の X に関する近似的弾性値とされるものである．今この関係を利用して目的変量 Y の平均を推定することを考えると，定理 1.1 により $e^{\varphi(X)}=kX^{\beta}$ ($k>0$, 常数) に比例した確率抽出を行えばよいことになる．

ところで定理 1 は X を k 次元ベクトル値変数 $\boldsymbol{X}$ とした場合にも成立することが容易に証明される.[註3)] 従って一般的には $e^{\varphi(\boldsymbol{X})}=C\prod_{i=1} X_i^{\beta_i}$ 比例確率抽出法を最適抽出法とすることが出来る．具体的に例えば付加価置量 Y の平均値又は総額を推定する場合，コップ・ダグラス型の生産関数を予想して，その要素を補助変量 $\boldsymbol{X}$ とするならば，企業を $e^{\varphi(\boldsymbol{X})}$ に比例した確率で抽出するのが最適ということになる．このことは最適性は対象集団の内部構造に則した方式をとることによって得られるといえそうである．実際に定理 1.1 は二変量ジブラ分布が最もよい適用例を与えるのであり，パレート分布に対しても可成り有効であり，一般に所得分布に対して有効性をもつと考えられるのである．因にジブラ分布では

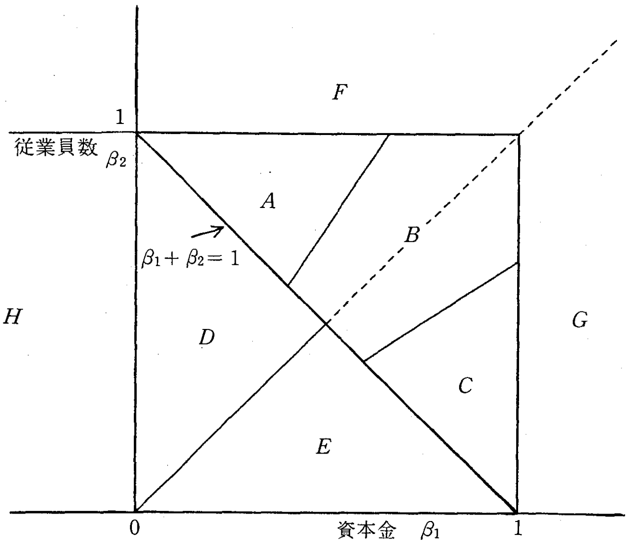
(1.6)
$$C^{(0)}=\frac{1}{\sqrt{n}}\sqrt{\exp\{(1-\rho_{xy}^2)\sigma_x^2\}-1}$$

となる．ここで ρ_{xy}, σ_x^2 はそれぞれ母集団の対数相関係数及び対数分散を表わすものである．勿論実際には (1.5) の成立は理想状態なのであり，パレート分布を前提とすると ζ は X と一定の相関を示すのである．このことは次章に示すより一般的な方法を導く一つの導因となっている．

扨，分析を更に進めると，$\sum\beta_i$ は 1 に近い場合が多く又単一の補助変量に比例した抽出確率によっても或程度効果がある事，更に ζ が X と相関関係にある場合は，或程度の系統抽出の導入が必要となる事情が予想されるのである．この系統抽出の導入は，再び有意選出法への接近を思わせるが，それは実際に次章にのべる形式に近いものである．

更に無視出来ないことは，(1.5) 式が可成りよい適合性を示す場合は，弾性値が抽出確率を決定する基本的要素となることである．つまり，このことを産業特性による地域層化や企業層化に対しても弾性値を基準にすべきことを示している．これによって従来みられなかった層化についての一つの客観的基準が与えられることになり，その為に対象の構造への接近が求められるのである．

実際にそのような弾性値の構造を企業の産業別データによって吟味してみよう．その為にま



第1.1図 弾性値領域の設定

ず企業を対象として，対数線形回帰による有形固定資産の資本金及び従業員数に関する弾性値 β_1, β_2 を求め，第1.1図のような領域分割に基づいて産業を層化してみよう．その結果，資産の種類に応じて第1.1, 1.2, 1.3の各表の分類が得られる．[註4)]

実際に層別した結果が，調査結果に効果をもたらす為には，層化に用いた資料が調査時点における状態をある程度反映しているものでなければならない．その為には，弾性値の年次別・時期別変化がどの程度のものであるかを知る必要がある．第1.2図はその為の検討資料である．

又観点を変えて各資産・資本の総資産に関する弾性値を業種別に示したのが，第1.3図であり，業種別年次変化を示したのが第1.4図である．

此等の諸結果は，基本的な弾性値を注意深く選択するならば層化の基準として適用するのに充分適しているように思われる．第1.5図は更に此等の弾性値計算の根拠を与える(1.5)式の関係が，どの程度データの上で保障されるかを示したものであるが，この産業は比較的適合度の低い例とみてよい．我々は次の段階として此の図に示されるような回帰残差についても更に検討すべきであると考えている．

以上のように回帰関係を通して母集団の構造に立入ることは，抽出確率や層化の合理的な基準を与えることになり，任意抽出論を一歩前進させることになる．然しそれを実践に移す場合，例えば定理1.1の(1.3)式で与えられる $\varphi(X)$ や(1.5)式の係数 β は本来は事後的に得られるものであるから，それを事前に如何に予測するかが問われねばならない．又此等の回帰関数に対して残差 ξ が独立であるという条件も，例えばマルディアのパレート多変量分布等に対しては成り立たないという事情を考慮せねばならない．

又，一般的に言って任意標本抽出論における等確率抽出法から不等確率抽出法への推移は，任意抽出法の種類を増大し，その適用範囲を拡大したが，同時にそれは抽出法の選択の問題を生ずることになった．Optimalな選択が問われることになり，標本調査論は，実務上の指導よりも一層理論的な課題として研究室や教室に持ち込まれることになった．そのことは，統一理論の発生と進歩を促したが，その結論は必ずしも任意抽出法に対して肯定的なものばかりではない．例えばGodambe (1965) はサンプリングの統一理論に基づいて「任意のデザイン d に対し，母総計 τ に対するすべての線形不偏推定量のクラス $L(d)$ の中で，一般には一様最小分散

第 1.1 表 （有形固定資産―土地）の弾性値による分類

A：繊維(BDA)，衣服($AFAB$)，木材(HA)，一般機械，電機，輸機，精機(BA)，卸売，小売(BA)，
他サ$_{\triangle}$($BAHA$)
B：建設，食料，衣服，パルプ$_{\triangle}$，金属，石油・石炭(BD)，他製，放送
C：化学，窯業，鉄鋼(CB)，不動産，非鉄(CE)，陸運，水運，ガス(BC)，他運(ED)
D：船舶(DED)，事サ$_{\triangle}$(ED)，旅・宿
E：映・娯
F：鉱業
G：電気
H：出版$_{\triangle}$(HDH)
I：個サ

第 1.2 表　棚卸資産の弾性値による分類

A：鉱業(BAD)，パルプ，窯業，非鉄，金属，一般機械，精機，化製，卸売，小売，旅・宿，映・娯
B：建設$_{\triangle}$，化学，石油・石炭，鉄鋼，輸機，事サ$_{\triangle}$(EB)
C：食料，繊維，水運
D：衣服$_{\triangle}$(BD)，他運$_{\triangle}$
E：
F：木材，不動産(HF)，ガス(FAF)，他サ$_{\triangle}$(AHF)
G：陸運，電気(CG)，個サ$_{\triangle}$
H：船舶(HFH)
I：

第 1.3 表 （有形固定資産―土地）＋棚卸資産の弾性値による分類

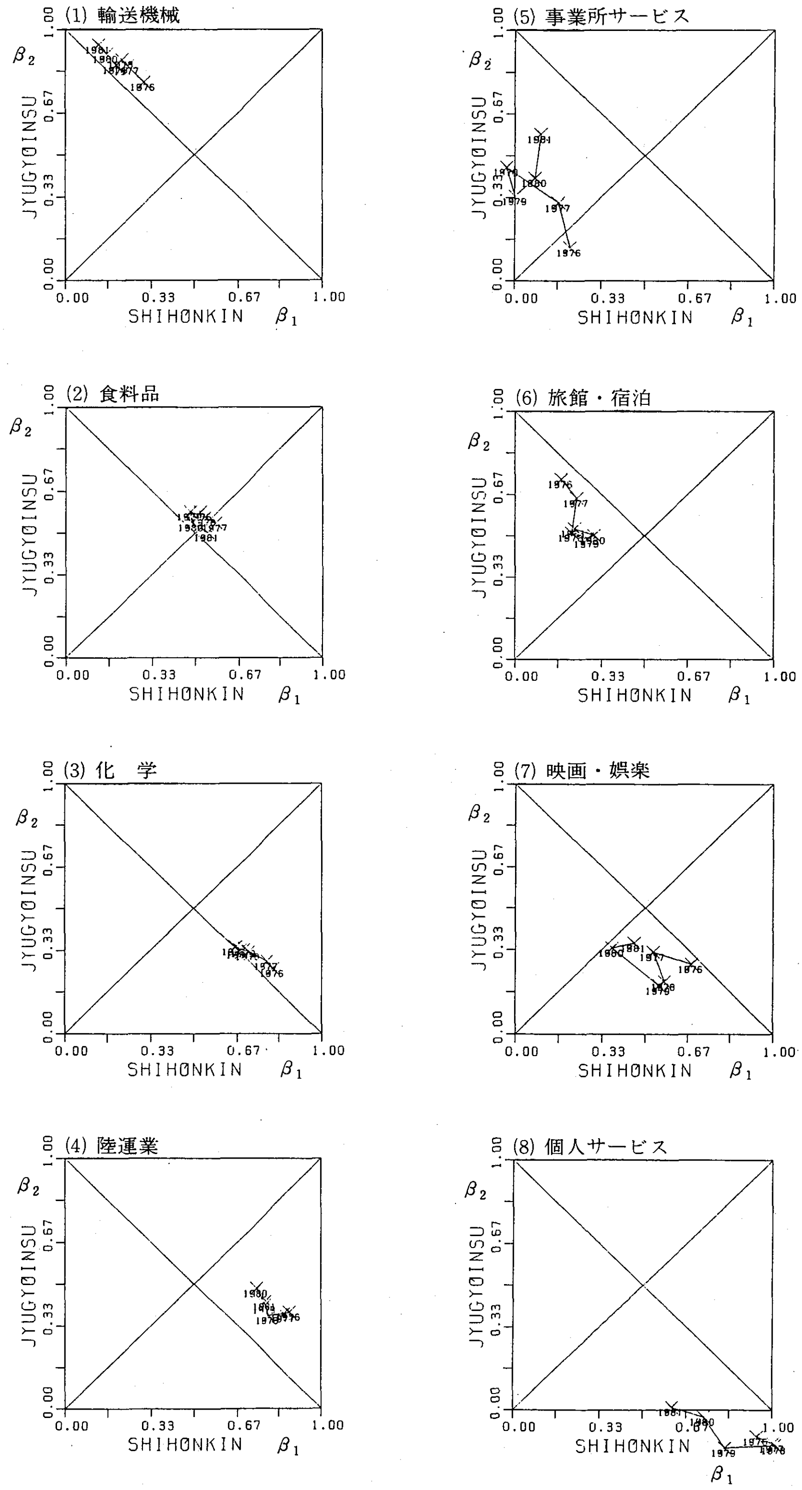
A：木材，パルプ，出版$_{\triangle}$(ABA)，金属，一般機械，電機，輸機，精機，船舶(ADA)，卸売，小売
B：建設，食料，繊維，石油・石炭(BA)，窯業(CB)，鉄鋼，非鉄，他製，不動産，放送
C：化学，陸運，水運，他運，ガス
D：衣服(DAD)，旅・宿
E：事サ$_{\triangle}$(DE)，映・娯
F：鉱業
G：電気
H：他サ$_{\triangle}$($AFHD$)
I：個サ

(UMV)不偏推定量は存在しない．それが存在するのは次のようなクラスD(ユニクラスター・デザイン）に限られる．

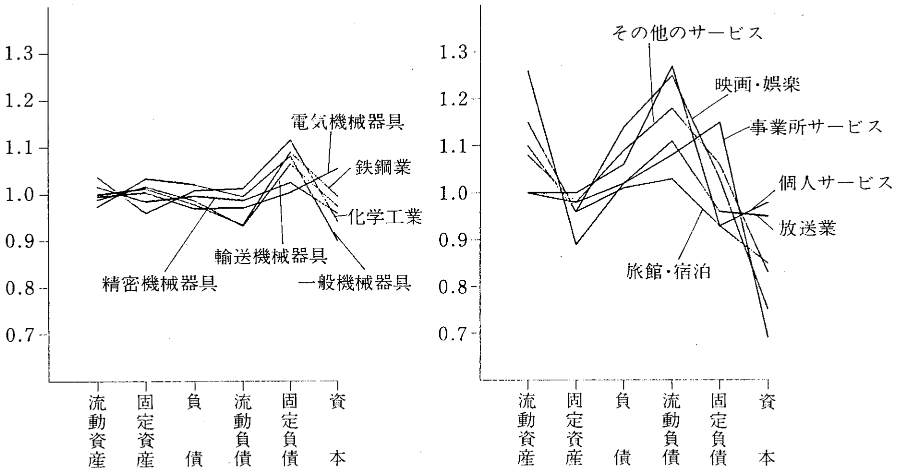
$P(S_1)>0$ かつ $P(S_2)>0$ であるとき，これら2組のサンプルは等価 $S_1\sim S_2$（有効サンプルが一致）か又は互いに素 $(S_1\cap S_2=\phi)$ である.」(多賀 (1976) p. 222) を見出している．

かくして最小分散を optimal とする立場は他の条件としてモデルを導入することになり，それと異なる立場は optimal の基準を他に求めることになった．

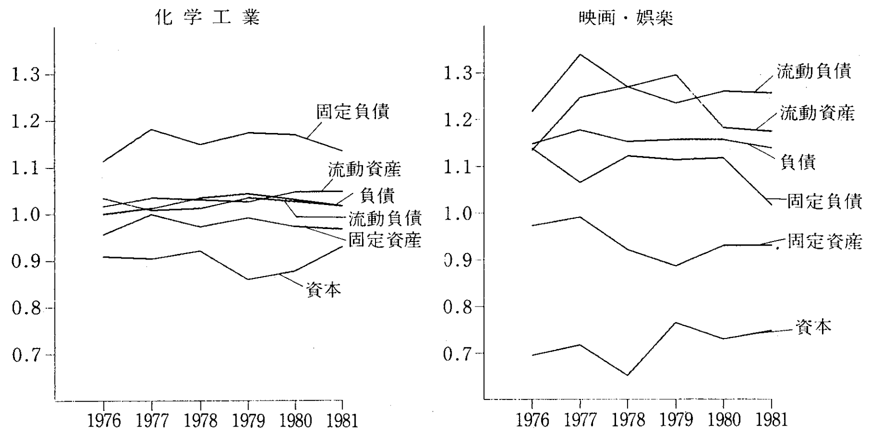
前者の立場では，例えば Royall (1970a) は linear model depend を前提として任意抽出法に対して否定的な結論を導いている．すなわち有限母集団に於いて補助変量 x_k を凡て既知としたとき，有限母集団の母平均 $\bar{Y}$ の不偏推定量のうちで MSE を最小にする最適な推定量と最



第1.2図　(有形固定資産—土地) の弾性値の産業別年次変化



第 1.3 図　総資産に関する弾性値の業種別変動



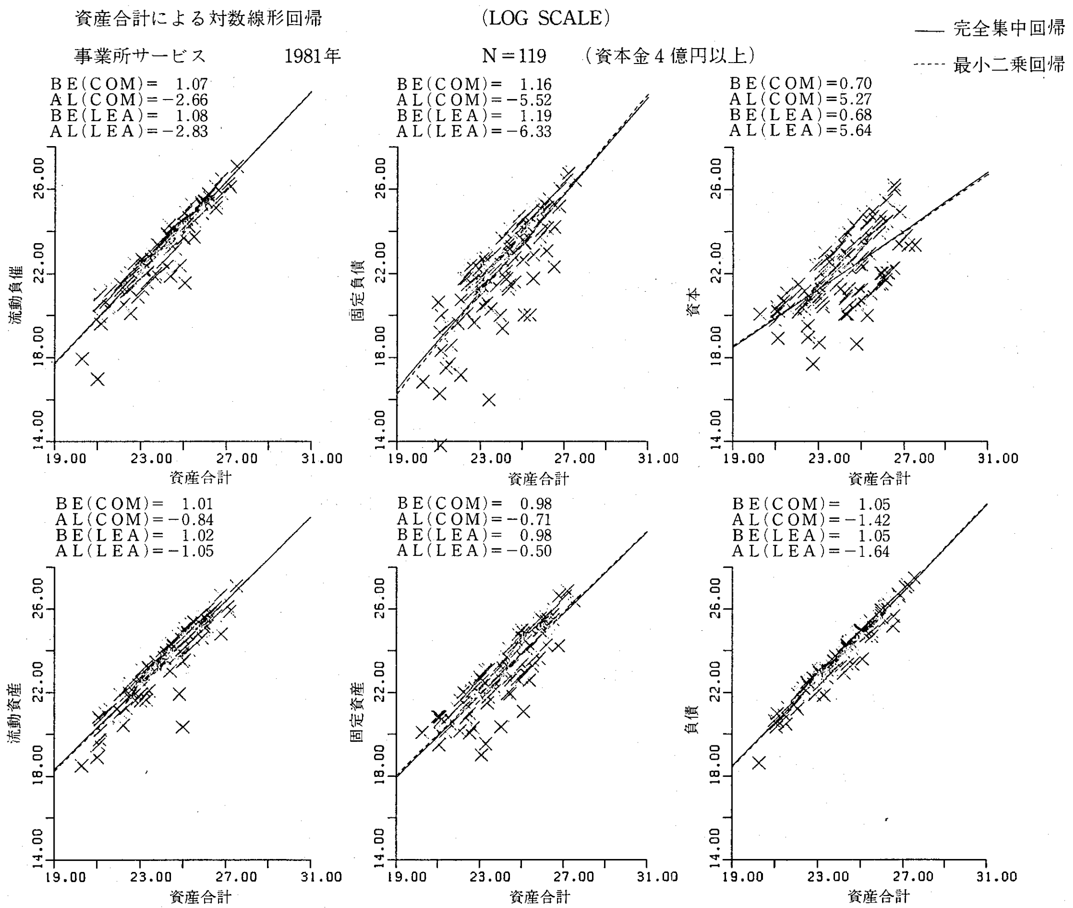
第 1.4 図　総資産に関する弾性値の年次変化（完全集中回帰）

適なサンプリング・デザイン s^* についての sampling plan p^* は

(1.7)
$$p^*(s)=\begin{cases}1 & s=s^*\\ 0 & s\neq s^*\end{cases}$$

であることを見出した．その結果彼はこの最適標本を non random であるという意味で有意選出 purposive sampling として促えた．一方，optimal の基準を他に求める方向は，多くの既成の数理統計学的諸観点（充分性や完備性）に基づいて検討されており，それは前掲書 多賀(1976)の第 9 章の後半の主題として詳細にのべられている．そしてその結論として著者は次のように述べているのである．

「以上の史的概観を通じてみると，不等確率サンプリングは当初より良いサンプリング・デザインを見出すことを意図し，比例確率サンプリングなど実際に有効な方法を生み出したが，その後サンプリング統一理論への適用に興味の焦点がうつっていった．しかしながら，あるサンプリング・デザインに対して最良な線形推定量を求めると，その係数は母集団パラメータに依



第1.5図 弾力性構造の信頼性

存したものとなり，それは推定量とはいえない．これは母集団について『どれだけでも情報がえられる』という現実を無視した前提に立って議論をしているからに外ならない．層別サンプリングや重ねサンプリングなどは，まさしく現実にえられる情報の範囲で可能なサンプリング・デザインを与える方法であり，今後はそのように入手可能な情報の利用に限定した範囲で一般的な議論を展開すべきであろう.」(第9章の末尾 p. 224)

又第2節の末尾では，

「以上の事柄を要約すると，サンプリング方式および推定法においては，母集団の個々の要素について事前に何の情報もえられない場合には非復元単純サンプリングを行って普通の推定を行うしかないが，何等かの有力な事前情報が入手できる場合には，それを有効に生かせるようなサンプリング方式（デザイン）と推定法の選定に工夫をこらすべきであるといえよう．この点に関する理論と方法の研究は十分行われているとはいい難い現状であり，将来における発展が期待される.」(p. 235)

とのべている．

此等の結果や結論は任意抽出法の将来の可能性を否定するものではないが，それは同時にかってネイマンにより否定された有意選出法の再検討を行う道が開かれたことも否めないのである．

かくして我々は，統一理論に基づく此等の研究成果を他山の石として，次章に於いて独自の立場で有意選出法を一般化して検討することにしよう．それは最早抽象的な一般論としてではなく，180°転回して具体的実践的な提案を通して行われる．

註 1) 特に層化の為の最適分点の決定については次のような統計数理研究所における一連の諸論文を挙げることができる．すなわち Hayashi, C., F. Maruyama and M. D. Ishida (1951), 林知己夫，多賀保志，高倉節子 (1955)，林知己夫，丸山文行 (1949)，増山元三郎 (1950), Taga, Y. (1967, 1971), Wakimoto, K. (1971).

他方，最適変動抽出確率を求めたものとしては，後掲の田口（1975, 1976a, 1976b, 1978）を挙げることが出来る．

註 2) 定義 1.1 及び定理 1.1 について後掲 田口（1976a）に詳しい．

註 3) 前掲 田口（1976a）参照．

註 4) 以下の諸資料は法人統計年報（1976-1981）によるものである．特に第 1.1～3 表における△印は最小二乗回帰による弾性値と集中回帰による弾性値の差の大きい産業を示している．又（ ）内の記号は年変化によって複数の領域にまたがる場合を示している．

2. 有意標本の選出法と推論*

前章に示したように変動確率抽出法の最適化の試みは全体及び部分の精度の向上と共に，層化についての一つの客観的規準を与えるものであった．それは又全体として母集団構造の解析が標本調査法の改良に役立つことを示した．然し乍ら最適化理論を成立させる条件は，広く非線形回帰構造を前提とするものではあったが，なお残差のコントロールに対する独立性等の制約を伴うものであった．又有限母集団からの非復元抽出は統一理論によれば model depend にせよ non model depend にせよ最適性の追求は non random な標本に導くものであった．

序論に触れた方法論史的考察によれば，我々は此の段階において獲得した任意抽出法の成果を有意選出法の立場で吸収し，更に高度の標本調査法を形成し得る筈である．我々は本章で実際にその実現可能性を検討することにしよう．その理論的根拠は，第 2.1 図に示す回帰適合曲線の構成によって与えられる．然しここでは此の曲線概念に触れる前に，まず予想される有意選出法について具体的な解説を加えることにする．

2.1. 有意選出法の定式化

ジーニまでの有意選出法は，多くは単純な例示に止まり，定式化するまでには至っていない．一方ロイヤルは統一理論の立場に立って，しばしば有意選出法を口にするが，対象集団のもつべき条件も，具体的な選出手続きも示していない．従って我々は有意選出法の定式化から出発せねばならない．

1. 目的

y を未知の目的変量（調査特性），x を既知の補助変量（コントロール）として y の平均値 μ_y を有意選出標本によって推定する方式を考える．

2. 選出方法

標本選出の原理は y の x による近似的回帰関数値 $\varphi(x)$ で加重した x の順位に従った系統選出である．又その選出の始点は，この加重選出間隔の中間を占める個体である．すなわち $\varphi(x)$ を x の順位に従って累計し，その各累計額を累計総額の $1/2k$ で等分割した時，奇数番目の分割点を占める個体を標本とする．

* 本章以下は昭和 61 年度日本統計学会における筆者の報告「有意選出法について」の詳論である．

但しもし回帰関数が上記の選出間隔すなわち（$\varphi(x)$の累計総額)$/k$を超える個体があれば予めそれを全数調査対象として除外し，残部について上記の手続きによって標本を選出する．

又もし補助変量として x_i をとる個体が標本に含まれる場合は，原則上 x_i を有する凡ての個体を標本とする．従って部分的にも全数調査部分を含むことになり，選出時における同一 x 値を占める個体間の順位は考慮する必要がなくなる．

3．選出の条件

扨，$R(x)$ を y の x に関する回帰関数，すなわち $\mu_{y|x}$ をコントロール x をとるグループの y の平均値（y の x に関する条件付平均）とした時，

$$R(x)=\mu_{y|x} \tag{2.1.1}$$

とする．又 $\tilde{R}(x)$ を y の x に関する近似的又は予想回帰関数とする．その場合 μ_R 及び $\mu_{\tilde{R}}$ をそれぞれ $R(x)$ 及び $\tilde{R}(x)$ の全体平均値とすれば

$$\begin{aligned}\mu_R&=\mu_y\\ \mu_{\tilde{R}}&\neq\mu_y\end{aligned} \tag{2.1.2}$$

である．

この時仮定 H_1 及び H_2 として

H_1: $R(x)$ 及び $\tilde{R}(x)$ は x の連続な正値関数である．

H_2: x に関する母集団分布は連続な密度関数 $f(x)$ によってよく適合される．

このとき以下に述べる有意選出法及び推定法の基礎となるのは x をパラメータとする次の曲線（仮に回帰適合曲線と呼ぶ）$\boldsymbol{C}=(V,W)$ である．即ち

$$\begin{aligned}V(x)&=\int_0^x \tilde{R}(\xi)f(\xi)d\xi/\mu_{\tilde{R}}\\ W(x)&=\int_0^x R(\xi)f(\xi)d\xi/\mu_y\end{aligned} \tag{2.1.3}$$

である．従って横軸は既知量，縦軸は未知量となる．このとき W は V の一価関数と考えられるから，それを

$$W=\phi(V) \tag{2.1.4}$$

とする．

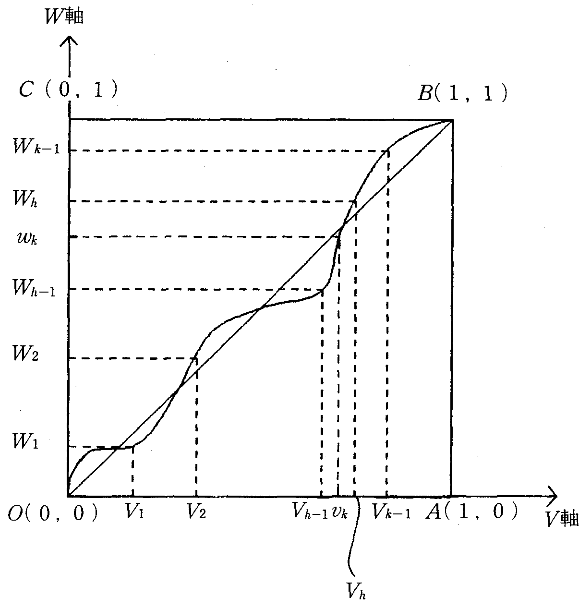
この曲線はローレンツ曲線に類似しており，仮説 H_1, H_2 のもとで

$$0\leqq V\leqq 1,\quad 0\leqq W\leqq 1 \tag{2.1.5}$$

$$\phi(0)=0,\quad \phi(1)=1 \tag{2.1.6}$$

となる．従って $\boldsymbol{C}$ は第2.1図に示すような点 O と点 B を結ぶ単調に増加する曲線である．但しローレンツ曲線と異なり，V 軸に対して常に凸又は凹ではない．然し極めて興味のあることはこの凹凸は後出 (2.1.33) 式で示されるように $R(x)$ の局所弾力性が $\tilde{R}(x)$ のそれよりも大なるとき V 軸に対しては凸，小なるときは凹となる．又両弾力性が一致する点が変曲点となる．

又特に $\tilde{R}(x)$ が $R(x)$ と一致するときは直線 OB となる．これを適合基準線ということにする．実際にはこのような場合は考えられない．[註1)] 直線 OB はローレンツ曲線の均等線に対応するものであるから，この曲線に関しても集中面積やジーニ集中度係数にあたる特性量が考えられる．すなわち回帰適合曲線と基準線 OB とで囲まれる面積を適合面積 A とすれば



第2.1図　回帰適合曲線

(2.1.7)
$$A=\frac{\varDelta_{R\cdot\tilde{R}}}{4\mu_R\mu_{\tilde{R}}}$$
$$\varDelta_{R\cdot\tilde{R}}=\int_0^\infty\int_0^\infty \operatorname{sgn}(x_2-x_1)\begin{vmatrix}\tilde{R}(x_1) & \tilde{R}(x_2)\\ R(x_1) & R(x_2)\end{vmatrix} f(x_1)\,f(x_2)dx_1dx_2$$

となる．但し曲線が規準線の下部にある部分の面積を正，上部にある部分の面積を負とする．従って $F=2A$ を回帰適合度とすれば一般に

(2.1.8)
$$-1\leqq F\leqq 1$$

が成立する．特に x の凡ての値に対して $R(x)/\tilde{R}(x)$ が定数の場合は $F=0$ となる．従ってもし $R(x)$ と $\tilde{R}(x)$ が常に定弾力性をもつとすれば，その両者が等しい場合 F は0となる．$\varDelta_{R\cdot\tilde{R}}$ を回帰適合差とよぼう．

この回帰適合曲線をもとにして，以下選出の方法と推定の方式を具体的に示すことが出来る．

4．選出の方法

標本が示すコントロール x の種類を k 個 $x_1, x_2, \cdots, x_k$ ときめる．但し母集団における $\tilde{R}(x)$ の総額を $T_{\tilde{R}}$ としたとき $\tilde{R}(x)\geqq T_{\tilde{R}}/k$ となるような個体は凡て全数調査対象として予め除外しておく．

扨，$V(x)$ の取る値を k 個の等区間に分割する．この際，区間 OA の分割点

(2.1.9)
$$V_h\equiv V(x_h);\ h=1, 2, \cdots, k-1$$

で分けられた各小区間の巾が $\dfrac{1}{k}$ になるように x_k を決める．特に

(2.1.10)
$$V_0=0,\ \ V_k=1$$

とする．このとき h 番目の区間 I_h を $(V_{h-1},\ V_h]$ と表すことができ，

(2.1.11)
$$V_h-V_{h-1}=1/k;\ h=1, 2, \cdots, k$$

とする．更に第1図に示すように此等の分割点に対応して

(2.1.12) $$W_h = W(x_h);\ h=0, 1, 2, \cdots, k$$

とする．このとき明らかに

(2.1.13) $$W_0=0,\ W_k=1$$

となる．

さて，I_h に収まる部分母集団 π_h において $V(x)$ の区間中間値を占める個体が示すコントロールの値を $\hat{x}_h$ とすれば

(2.1.14) $$v_h \equiv V(\hat{x}_h) = (V_{h-1}+V_h)/2$$

である．v_h に対応して

(2.1.15) $$w_h \equiv W(\hat{x}_h)$$

が決まる．もし次の仮説 H_3

H_3: $\tilde{R}(x)$ は x の単調関数である．

が成立するとき各 $\tilde{R}(\hat{x}_h)$ は π_h における $\tilde{R}(x)$ に関するメディアールとなる．然しこの仮説は特に必要でない．[註2)]

一般的にみて $\hat{x}_h$ をとる個体は1個ではない．いまその全体を全て標本に加え，この部分における y についての平均値を考えるとそれが $R(\hat{x}_h)$ である．又この場合 $\hat{x}_h$ をもつ個体の配列は問題とならない．

具体的にサイズ N の有限母集団 π を対象として H_2 が成立するものとする．このときコントロールが $\hat{x}_h$ の近傍に属する個体のクラスを c_h とすると c_h の標本数 n_h は近傍の区間の巾を Δ_h とすれば

(2.1.16) $$n_h \fallingdotseq Nf(\hat{x}_h)\Delta_h$$

となる．従って全標本数 n は

(2.1.17) $$n=\sum_{h=1}^{k} n_h \fallingdotseq N \sum_{h=1}^{k} f(\hat{x}_h)\Delta_h$$

である．ここで Δ_h は例えば

(2.1.18) $$V\left(x_h+\frac{\Delta_h}{2}\right)-V\left(x_h-\frac{\Delta_h}{2}\right)=\frac{1}{2kl}$$

となるようにすればよい．この時各小区間で選出される標本の $\tilde{R}(x)$ に関するシエアは一率に $\frac{100}{2kl}\%$ となる．

5．推定の方式

(2.1.13), (2.1.14), (2.1.15) の諸式から容易に

(2.1.19) $$V_h=v_h+\frac{1}{2k},\ V_{h-1}=v_h-\frac{1}{2k}$$

が得られる．従って又 (2.1.4) 式から

(2.1.20) $$W_h=\phi\left(v_h+\frac{1}{2k}\right),\ W_{h-1}=\phi\left(v_h-\frac{1}{2k}\right)$$

となる．今仮説 H_4 として

H_4: $\phi(V)$ は Taylar 展開可能である．

としよう．このとき H_4 の下で

$$W_h-W_{h-1}=\phi\left(v_h+\frac{1}{2k}\right)-\phi\left(v_h-\frac{1}{2k}\right) \tag{2.1.21}$$
$$=2\left\{\frac{1}{2k}\phi'(v_k)+\frac{1}{3!}(2k)^3\phi'''(v_k)\right\}+O\left(\frac{1}{k^5}\right)$$

となる．ところで

$$\sum_{h=1}^{k}(W_h-W_{h-1})=1 \tag{2.1.22}$$

であるから

$$2\sum_{h=1}^{k}\left\{\frac{1}{2k}\phi'(v_k)+\frac{1}{3!}\frac{1}{(2k)^3}\phi'''(v_k)\right\}+O\left(\frac{1}{k^4}\right)=1 \tag{2.1.23}$$

となる．更に (2.1.3), (2.1.4) 式によると

$$\phi'(V)=\frac{W'(x)}{V'(x)}=\frac{R(x)}{\tilde{R}(x)}\frac{\mu_{R\tilde{R}}}{\mu_y} \tag{2.1.24}$$

であるから (2.1.23) 式に代入すると

$$\mu_y=2\sum_{h=1}^{k}\left\{\frac{1}{2k}\frac{R(\hat{x}_h)}{\tilde{R}(\hat{x}_h)}\mu_{\tilde{R}}+\frac{\mu_y}{3!(2k)^3}\phi'''(v_k)\right\}+O\left(\frac{1}{k^4}\right) \tag{2.1.25}$$

そこで μ_y の第 1 推定量を $\bar{Y}_1$ として

$$\bar{Y}_1=\frac{1}{k}\sum_{h=1}^{k}\bar{Y}_{1h}=\frac{\mu_{\tilde{R}}}{k}\sum_{h=1}^{k}\frac{R(\hat{x}_h)}{\tilde{R}(\hat{x}_h)} \tag{2.1.26}$$

を考察することにする．ここで $R(\hat{x}_h)$ は調査によってきまる c_h に属するすべての個体の特性 Y に関する平均値 $\hat{y}_h$ で表わす．すなわち

$$\hat{y}_h=\frac{1}{n_h}\sum_{i=1}^{n_h}y_{h_i}\fallingdotseq R(\hat{x}_h) \tag{2.1.27}$$

とする．但し y_{h_i} は c_h クラスの各個体の調査特性値である．

扨，(2.1.26) 式の $\bar{Y}_1$ を μ_y の推定量とした場合それは (2.1.27) 式により標本測定量 y_{h_i} の線形推定量となる．又この時 (2.1.25) 式の第 2 項は $\bar{Y}_1$ の推定誤差或いはその補正項と考えられるが，それを $e_{\bar{Y}}$ とすると

$$e_{\bar{Y}}=\frac{\mu_y}{4\cdot 3!k^3}\sum_{h=1}^{k}\phi'''(v_k) \tag{2.1.28}$$

とかける．[註3)] 次にこれを改めて吟味する．

6.　$\tilde{Y}$ についての推定精度

(2.1.24) 式を V に関して対数微分すると容易に

$$\frac{\phi''(V)}{\phi'(V)}=\left\{\frac{R'(x)}{R(x)}-\frac{\tilde{R}'(x)}{\tilde{R}(x)}\right)/v(x) \tag{2.1.29}$$

が得られる．ここで $v(x)$ は x 値クラスの母集団における $\tilde{R}(x)$ に関する限界占有率

(2.1.30)
$$v(x)=V'(x)=\frac{\tilde{R}(x)}{\mu_{\tilde{R}}}f(x)$$

を表わす.

今$R(x)$及び$\tilde{R}(x)$のxに関する弾性値をそれぞれ

(2.1.31)
$$\eta_R(x)=\frac{d\log R(x)}{d\log x}=\frac{xR'(x)}{R(x)}$$
$$\eta_{\tilde{R}}(x)=\frac{d\log\tilde{R}(x)}{d\log x}=\frac{x\tilde{R}'(x)}{\tilde{R}(x)}$$

としてその差を

(2.1.32)
$$\delta_\eta(x)=\eta_R(x)-\eta_{\tilde{R}}(x)$$

とする．この時(2.1.29)式は

(2.1.33)
$$\frac{\phi''(V)}{\phi'(V)}=\frac{\delta_\eta(x)}{xv(x)}$$

となる．これを更にVに関して対数微分すると

(2.1.34)
$$\frac{\phi'''(V)}{\phi''(V)}-\frac{\phi''(V)}{\phi'(V)}=\left\{\frac{\delta'_\eta(x)}{\delta_\eta(x)}-\frac{1}{x}-\frac{v'(x)}{v(x)}\right\}/v(x)$$

となり，(2.1.33), (2.1.34)式から結局

(2.1.35)
$$\frac{\phi'''(V)}{\phi'(V)}=\frac{\delta_\eta(x)}{xv^2(x)}\left\{\frac{\delta_\eta(x)}{x}+\frac{\delta'_\eta(x)}{\delta_\eta(x)}-\frac{1}{x}-\frac{v'(x)}{v(x)}\right\}$$

が得られる．今$|\delta_\eta(x)|$及び$v(x)$のxに関する弾性値をそれぞれ

(2.1.36)
$$\delta_\eta^{(\cdot)}(x)=\frac{d\log\delta_\eta(x)}{d\log x}=\frac{x\delta'_\eta(x)}{\delta_\eta(x)}$$
$$v^{(\cdot)}(x)=\frac{d\log v(x)}{d\log x}=\frac{xv'(x)}{v(x)}$$

とすれば$\phi'''(V)$は更に

(2.1.37)
$$\phi'''(V)=\frac{\phi'(V)\delta_\eta(x)}{x^2v^2(x)}\{\delta_\eta(x)+\delta_\eta^{(\cdot)}(x)-v^{(\cdot)}(x)-1\}$$

となる．

結局(2.1.24), (2.1.28), (2.1.37)の各式により

(2.1.38)
$$e_{\bar{Y}_1}=\frac{\mu_{\tilde{R}}}{4\cdot 3!k^3}\sum_{h=1}^{k}\frac{R(\hat{x}_h)}{\tilde{R}(\hat{x}_h)}\frac{\delta_\eta(\hat{x}_h)}{\hat{x}_h^2v^2(\hat{x}_h)}\{\delta_\eta(\hat{x}_h)+\delta_\eta^{(\cdot)}(\hat{x}_h)-v^{(\cdot)}(\hat{x}_h)-1\}$$

と書くことが出来る．

従ってkが比較的大きく，サンプルの$\tilde{R}(x)$に関する占有率が高く，且つyのxに関する近似的回帰関数がよいもので弾性値の差が小さければ誤差は小となるのである．

ところで(2.1.26)式に於いて，$\hat{x}_h$をとる個体が複数n_h個存在すれば，凡て標本に加えるものとした．(局所的全数調査)．又，4.選出の方法で述べたように高額の$\tilde{R}(x)$をもつ個体も全数調査の対象となりうるものであった．従って標本数は厳密には事前に確定し難いといえる．

然し乍ら経済調査のような場合n_hは二，三の実験を重ねることによって容易にその範囲が経験的に確定されるし，又金額評識量は適当な切捨単位の設定によってn_hを調整することが出来る．

更に本稿では取上げないが，複数のコントロールを用いた多重及至多段選出の場合には，n_h は更に制限された値をとることになる．

2.2. 有意選出法の構造依存性

前節 (2.1.26) 式で示される有意選出法に基づく母平均の第1推定方式は，比推定形式をとるから，見掛け上は $\tilde{R}(x)$ 比例確率抽出法における不偏推定量と同一である．然しそれを補正する精度 $c_{\bar{Y}}$ は，この有意選出法に特有のもので，(2.1.38) 式によって理解されるように，弾性値を通して，母集団の回帰構造とその近似形式に大きく依存する．以下そのことを具体的に検討してみよう．

〈1〉 $R(x)$ が x の1次同時式となる母集団構造

この場合，一般的に

(2.2.1) $$R(x)=\beta x$$

であるから

(2.2.2) $$\eta_R(x)=1,\ \delta_\eta(x)=1-\eta_{\tilde{R}}$$

となる．従って (2.1.38) 式により近似的回帰関数 $\tilde{R}(x)$ として x の1次同時式

(2.2.3) $$\eta_{\tilde{R}}(x)=\tilde{\beta}x$$

を用いれば $e_{\bar{Y}_1}=0$ となり，第1推定方式は単純比推定形式は

(2.2.4) $$\bar{Y}_1=\frac{\mu_x}{k}\sum_{h=1}^{k}\frac{\hat{y}_h}{\hat{x}_h}$$

であり，その推定効果は高いものとなる．

〈2〉 $R(x)$ が x の線形回帰となる集団構造

この場合一般的に

(2.2.5) $$R(x)=\alpha+\beta x$$

と表わせるから，近似回帰関数として同様に次の線形関数 $\tilde{R}(x)$ を用いることにする．すなわち

(2.2.6) $$\tilde{R}(x)=\tilde{\alpha}+\tilde{\beta}x \qquad \text{従って}$$
$$\mu_{\tilde{R}}=\tilde{\alpha}+\tilde{\beta}\mu_x$$

このとき容易に

(2.2.7) $$\eta(x)=1-\frac{\alpha}{\alpha+\beta x}$$
$$\tilde{\eta}(x)=1-\frac{\tilde{\alpha}}{\tilde{\alpha}+\tilde{\beta}x}$$
$$v(x)=\frac{\tilde{\alpha}+\tilde{\beta}x}{\tilde{\alpha}+\tilde{\beta}\mu_x}f(x)$$

が得られるから

(2.2.8) $$\delta_\eta(x)=\frac{(\tilde{\alpha}\beta-\alpha\tilde{\beta})x}{(\alpha+\beta x)(\tilde{\alpha}+\tilde{\beta}x)},\ \delta^{(\prime)}(x)=-1+\frac{\alpha}{\alpha+\beta x}+\frac{\tilde{\alpha}}{\tilde{\alpha}+\tilde{\beta}x}$$

(2.2.9) $$v^{(\prime)}(x)=\tilde{\eta}(x)+f^{(\prime)}(x) \quad \text{但し} \quad f^{(\prime)}(x)=\frac{xf'(x)}{f(x)}$$

となる．従って

(2.2.10)
$$\frac{\tilde{\beta}}{\tilde{\alpha}}=\frac{\beta}{\alpha}$$

となるように近似的線形回帰が得られれば第1推定方式

(2.2.11)
$$\bar{Y}_1=\frac{\tilde{\alpha}+\tilde{\beta}\mu_x}{k}\sum_{h=1}^{k}\frac{\hat{y}_h}{\tilde{\alpha}+\tilde{\beta}\hat{x}_h}$$

は極めて有効なものとなる．

一方 (2.2.7), (2.2.8), (2.2.9) の諸式により

(2.2.12)
$$\begin{aligned}\delta_\eta(x)+\delta^{(\cdot)}(x)-v^{(\cdot)}(x)-1&=\frac{3\tilde{\alpha}}{\tilde{\alpha}+\tilde{\beta}x}-f^{(\cdot)}(x)-3\\&=-3\tilde{\eta}(x)-f^{(\cdot)}(x)\end{aligned}$$

となるから (2.1.38) 式から結局

(2.2.13)
$$e_{\bar{Y}_1}=\frac{(\tilde{\alpha}+\tilde{\beta}\mu_x)^3}{4\cdot 3!k^3}\sum_{h=1}^{k}\frac{\tilde{\alpha}\beta-\alpha\tilde{\beta}}{\hat{x}_h(\alpha+\beta\hat{x}_h)}\frac{\hat{y}_h}{(\tilde{\alpha}+\tilde{\beta}\hat{x}_h)^4f^2(\hat{x}_h)}\left\{\frac{3\tilde{\alpha}}{\tilde{\alpha}+\tilde{\beta}\hat{x}_h}-f^{(\cdot)}(\hat{x}_h)-3\right\}$$

となる．

更に今マルディアのパレート第1型分布

(2.2.14)
$$\begin{aligned}f(x,y;a,b,p,q)&=\frac{p(p+1)(ab)^{p+1}}{(bx-ay-ab)^{p+2}};\ k>a>0,\ y>b>0;\ p>1,\ q>1\\&=0;\qquad\qquad x\leqq a,\ y\leqq b;\ p>1,\ q>1\end{aligned}$$

の成立を仮定すればその周辺分布は

(2.2.15)
$$\begin{aligned}f_1(x;a,p)&=\frac{pa^p}{x^{p+1}};\ x>a>0,\ p>1\\&=0\quad;\ x\leqq a,\ p>1\end{aligned}$$

であるから

(2.2.16)
$$R(x)=b+\frac{bx}{ap},\ f^{(\cdot)}(x)=-(p+1)$$

が成立する．[註4] 従って (2.2.13) 式から容易に

(2.2.17)
$$e_{\bar{Y}_1}=\frac{(\tilde{\alpha}+\tilde{\beta}\mu_x)^3}{4\cdot 3!k^3}\sum_{h=1}^{k}\frac{\tilde{\alpha}-ap\tilde{\beta}}{\hat{x}_h(ap+\hat{x}_h)}\frac{\hat{y}_h}{(\tilde{\alpha}+\tilde{\beta}\hat{x}_h)^4}\frac{\hat{x}_h^{2(p+1)}}{p^2a^{2p}}\times\left\{p+\frac{3\tilde{\alpha}}{\tilde{\alpha}+\tilde{\beta}\hat{x}_h}-2\right\}$$

が得られる．従って此の場合単に

(2.2.18)
$$\frac{\tilde{\alpha}}{\tilde{\beta}}=ap$$

となるように決定すれば (2.2.11) の第1推定方式は有効なものとなる．条件 (2.2.18) は単に補助変量に関するパラメータのみを含むから，このような決定は事前に行うことが出来る．それは例えば後述する分布適合曲線の性質を利用して (2.2.35) 式の $F_{f\cdot\tilde{f}}$ を0にする条件でパラメータを決定することである．

パレート分布は一般に二次以上のモメントの存在を期待出来ないから，任意抽出法にとって少くとも小標本論的には厄介な対象であり，この種の分布をその現実性に拘らず従前より特異

分布として敬遠する傾向にあった．パレート第1種分布のもつ以上の性質は有意選出法の適用にとっては，この分布は極めてよい条件をもつことを示すものであり，この選出法はここに独自の領域を見出したといえる．

実際に，同じ線形回帰をもつにしても正規分布の場合 (2.2.10) 式の左辺は事前に得られるパラメータ μ_x, σ_x の他に事後的に得られる μ_y, σ_y 及び相関係数 ρ_{xy} を含むからこの選出法の適用にとってよい条件を備えているとはいえない．

又ガンマ分布は同様に線形回帰をもち (2.2.10) 式の右辺は (2.2.18) 式のように補助変量 x のパラメータ α, β を含み目的変量 y のパラメータは含まないが，x, y 間の対数相関係数 ρ を含むものである．その意味でこの方法の適用の面では正規分布とパレート分布の中間に位置するのである．

〈3〉 $R(x)$ が定弾性値をもつ非線形回帰関数を表わす集団構造

データ解析上で，経験的に普及している手法は，経済データ等に於いて定弾性値をもつ回帰関数 $R(x)$ を予想し，その近似的回帰関数 $\tilde{R}(x)$ として対数線形回帰を用いることである．

今 $R(x)$, $\tilde{R}(x)$ をそれぞれ

(2.2.19) $$R(x)=\gamma x^{\eta},\ \tilde{R}(x)=\tilde{\gamma}x^{\tilde{\eta}}$$

とする．この時明らかに

(2.2.20) $$\eta_R(x)=\eta,\ \eta_{\tilde{R}}(x)=\tilde{\eta}$$

従って (2.1.32), (2.1.36) 両式により

(2.2.21) $$\delta_\eta(x)=\eta-\tilde{\eta}=\delta\ （定数）$$
$$\delta^{(\prime)}=0$$

となる．結局この場合 (2.1.38) 式は (2.1.26) 式の $\bar{Y}_{1h}$ を用いると

(2.2.22) $$e_{\bar{Y}}=\frac{\delta}{4\cdot 3!k^3}\sum_{h=1}^{k}\frac{\{\delta-\tilde{\eta}-f^{(\prime)}(x)-1\}}{\hat{x}_h^2 v^2(\hat{x}_h)}\bar{Y}_{1h}$$

となる．ここで $f^{(\prime)}(x)$ は $f(x)$ の x に関する弾力性係数を表わす．(2.2.22) 式は一見 $\bar{Y}_1$ の誤差項 $e_{\bar{Y}_1}$ も $\hat{y}_1, \cdots, \hat{y}_k$ についての線形結合のように見えるが，実際上は δ, η の測定に $\hat{y}_1, \cdots, \hat{y}_k$ が用いられる為により複雑な形式をとることになる．然し (2.2.22) 式をみると誤差は y の x に関する弾力性係数 η を除き y の特性量に関係なく与えられるものである．(2.2.22) 式を更に仔細にみると，y の実際及び予想弾力性の他に x の分布の弾力性を通して分布構造に依存することが理解出来る．従って次に定弾力性をもつ主要な分布についてこれを具体的に検討してみよう．此等の分布は又代表的な所得分布として知られているものである．[註5)]

(i) マルディアのパレート第II型母集団の場合

この場合は，次節の (2.3.4) 式で示すように $R(x)$ は定弾力性をもつことが知られている．従って今 x に関する周辺分布を

(2.2.23) $$f_1(x)=\frac{pa^p}{x^{p+1}}\ ;\ 0<a\leqq x,\ p>1$$
$$=0\qquad ;\ x<a$$

とする．この時容易に

(2.2.24) $$f_1^{(\prime)}(x)=-(p+1)$$

が得られる．従って (2.2.22) 式は次節の (2.3.5), (2.3.6) 式を用いると

(2.2.25)
$$e_{\bar{Y}_1}=\frac{\delta(\delta-\tilde{\eta}+p)\,p^2a^{2\tilde{\eta}}}{4\cdot3!\,k^3(p-\tilde{\eta})^2}\sum_{h=1}^{k}\hat{x}_h^{2(p-\tilde{\eta})}\bar{Y}_{1h}$$
$$=\frac{\delta(\delta-\tilde{\eta}+p)p^3a^{3\tilde{\eta}}}{4\cdot3!\,k^3(p-\tilde{\eta})^3}\sum_{h=1}^{k}\hat{x}_h^{2p-3\tilde{\eta}}\hat{y}_h\,;\ \tilde{\eta}<p$$

と表される．従ってこの誤差項は x の巾関数で測定値 $\hat{y}_h$ を加重したものとなるが，δ は一般には調査後に $\hat{y}_h$ を用いて推定される．従って

(2.2.26)
$$\tilde{\eta}=\eta \quad 又は \quad \frac{p+\eta}{2}$$

とすれば $e_{\bar{Y}}$ は消滅するが，これは必ずしも設計時に見出すことは出来ない．今更に

(2.2.27)
$$\tilde{\eta}=\omega p\,;\ 0<\omega<1$$

とすれば η は次節 (2.3.5) 式で与えられるから，(2.2.26) の条件は

(2.2.28)
$$\omega=\frac{\alpha}{q-1+\alpha}<1 \quad 又は \quad \frac{q-1+2\alpha}{2(q-1+\alpha)}<1$$

となる．従って既存の資料により x，y 間の対数相関係数 α 及び y のパレート係数 q についての予想が得られるならば (2.2.28) 式に近い ω を得ることが出来る．一方 (2.2.27) 式を用いると (2.2.25) 式は

(2.2.29)
$$e_{\bar{Y}_1}=\frac{(\eta-\omega p)\{\eta+p(1-2\omega)\}a^{2\tilde{\eta}}}{4\cdot3!\,k^3(1-\omega)^3}\sum_{h=1}^{k}\hat{x}_h^{p(2-3\omega)}\hat{y}_h$$

と書ける．従って特に

(2.2.30)
$$\omega=\frac{2}{3}$$

とするならば

(2.2.31)
$$e_{\bar{Y}_1}=\frac{(3\eta-p)\,(3\eta-2p)}{8k^2}\bar{y}\,;\ \bar{y}=\frac{1}{k}\sum_{h=1}^{k}\hat{y}_h$$

と表わされる．

ところで (2.3.5) 式によると

(2.2.32)
$$0<\eta<p$$

であるから結局 $e_{\bar{Y}_1}$ は

(2.2.33)
$$-\frac{p^2\bar{y}}{2k^2}\leqq e_{\bar{Y}}<\frac{p^2\bar{y}}{4k^2}$$

となる．更に吟味すると

(2.2.34)
$$p=2\eta$$

で $\bar{Y}_1$ は最大の過大推定値となりその誤差は $\dfrac{p^2\bar{y}}{2k^2}$ である．又

(2.2.35)
$$\frac{p}{3}<\eta<\frac{2}{3}p$$

に於いて $\bar{Y}_1$ は過大推定量となるが $\dfrac{p}{3}$ 及び $\dfrac{2}{3}p$ に近づく程減少し遂に消滅する．更に

(2.2.36)
$$0<\eta<\frac{p}{3}\quad 又は\quad \frac{2}{3}p<\eta<p$$

に於いては $\bar{Y}_1$ は過小推定となり，それは0又は1に近づく程誤差が大となるが $\frac{p^2}{4k^2}$ を越える事はない．

この場合選出方法は補助変量 x のパレート係数 p のみによって決定される上に，第1推定方式 $\bar{Y}_1$ による誤差 $e_{\bar{Y}_1}$ はその推定に伴って(2.2.31)式から得られる．或は(2.2.28)式の関係から α 及び q の推定によっても得られる．又 η のよい推定量が得られない場合であっても，(2.2.33)式によってそれは充分コントロールされるのである．実際に調査に於いては複数の調査項目を対象とするから，特定の η に拘りなく此の式によって標本のサイズ k を決定出来ることの意義は大きい．

パレート母集団に関する一層詳細な解析は，次節で改めて行うことにする．その結果はよりよい精度を与えることになる．パレート母集団は，有意選出法と驚く程緊密な関係を保っている．例えばパレートは，$p=1.5$ を基準状態と考えていたようであるが[註6)]この場合は以上の結果によると $\tilde{\eta}=1$ つまり x の規模に比例した選出が実際的方法といえる．

(ii)　ジブラ分布の場合

x の密度関数を

(2.2.37)
$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma_x x}\exp\left\{-\frac{1}{2}\left(\frac{\log x-\mu_x}{\sigma_x}\right)^2\right\};\ x>0$$

とする．この時容易に

(2.2.38)
$$f^{(\prime)}(x)=-1-\frac{\log x-\mu_x}{\sigma_x^2}$$

が得られる．従って(2.2.19)式の $\tilde{R}(x)$ を用いると(2.2.38)式は

(2.2.39)
$$e_{\bar{Y}_1}=\frac{2\pi\sigma_x^2\exp(2\tilde{\eta}\mu_x+\tilde{\eta}^2\sigma_x^2)}{4\cdot 3!\,k^3}\delta\sum_{h=1}^{k}\left\{\delta-\tilde{\eta}+\frac{\log\hat{x}_h-\mu_x}{\sigma_x^2}\right\}\times\frac{1}{\hat{x}^{2\tilde{\eta}}}\exp\left(\frac{\log\hat{x}_h-\mu_x}{\sigma_x}\right)^2\bar{Y}_{1h}$$

と表現される．一見すると此の場合は，パレート分布の場合のように η について事前の情報をもたずによい推定結果を得ることは難しいように思われる．然し第4節で示すように δ について標本推定を同時に行えるのでそれによって $e_{\bar{Y}_1}$ が推定され，次節に述べるような第2推定方式によって推定精度を高めることが出来る．

然しそれにしてもパレート分布とジブラ分布は，変動確率抽出法によるか有意選出法によるかによって，その優劣の順位を変えるようである．つまり標本選出時の選択は，集団の構造に依存するのである．

〈4〉　$R(x)$ が変動する弾性値をもつ非線形回帰関数をもつ集団構造

此の場合に於いても，もし補助変量 x の z への適当な変換によって(2.2.19)で示されるような回帰関数 $R(z)$ が得られる場合は〈3〉と同様な考察を z をもとにして行うことが出来る．

然し一般的には(2.1.38)式は(2.1.30)，(2.1.32)，(2.1.36)の諸式に従って

(2.2.40)
$$e_{\bar{Y}_1}=\frac{\mu_{\tilde{R}}^{3}}{4\cdot 3!\,k^3}\sum_{h=1}^{k}\frac{R(\hat{x}_h)}{\tilde{R}^3(\hat{x}_h)}\frac{\delta_\eta(\hat{x}_h)}{\hat{x}_h^2 f^2(\hat{x}_h)}\times\{\delta_\eta(\hat{x}_h)-\eta_{\tilde{R}}(\hat{x}_h)+\delta_\eta^{(\prime)}(\hat{x}_h)-f^{(\prime)}(\hat{x}_h)-1\}$$

をみたす $\tilde{\eta}(x)$ を見出して適用することが第 1 推定方式を有効にすることになる．但しこの場合 $\delta(x)$, $\eta(x)$, $\tilde{\eta}(x)$ は一般に x の関数であるから，定弾性値の場合と異なり，より詳細な分析が必要となろう．例としてパレート，ジブラ両所得分布と並び代表的なものとされるアモローゾ（一般化ガンマ）分布を取り上げてみよう．[註7] この分布は例外的に線形回帰構造を示すものであるが，一般的には，次式に示すような変動弾力性構造をもっている．すなわち

(2.2.42)
$$f(x)=\frac{\alpha x^{\alpha-1}}{\Gamma\left(\dfrac{1}{h+1}\right)}\exp\left(-x^{\alpha h}\right);\quad x\geqq 0,\ \alpha,\ h>0$$

(2.2.43)
$$f^{(\prime)}(x)=\alpha-1-\alpha h x^{\alpha h}$$

(2.2.44)
$$R(x)=(1-\rho^h)^{\frac{1}{\beta h}}\frac{\Gamma\left(\dfrac{1}{h}+\dfrac{1}{\beta h}\right)}{\Gamma\left(\dfrac{1}{h}\right)}F\left(-\frac{1}{\beta h},\ \frac{1}{h};\ -\frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)$$

で得られる．

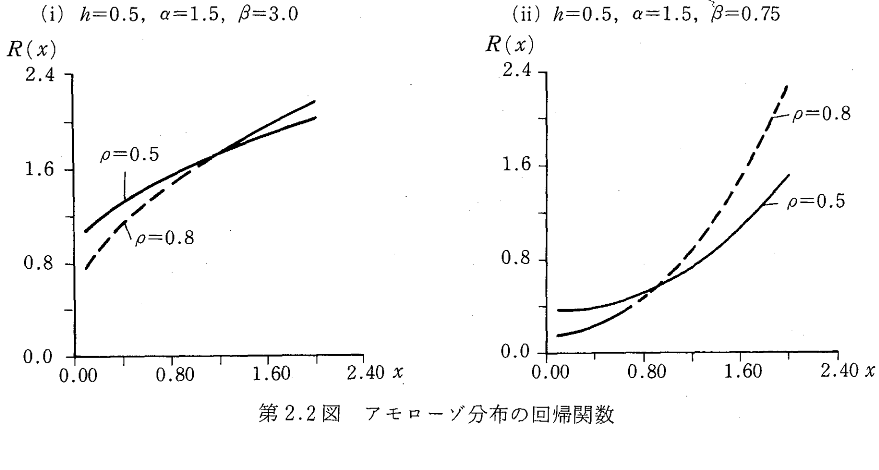
従って $\tilde{R}(x)$ としては $x^{\alpha h}$ に関する多項式を用いるのが望ましいことになる．（第 2.2 図参照）その場合分布のパラメータ及びパラメータ間の関係によっては，その多項式は単項式の巾関数でよく近似されることもあり得よう．

又 $R(x)$ の弾力性係数 $\eta(x)$ は

(2.2.45)
$$\eta(x)=-\frac{\alpha h\rho^h}{1-\rho^h}x^{\alpha h}+\frac{x\dfrac{d}{dx}F\left(\dfrac{1}{h}+\dfrac{1}{\beta h},\ \dfrac{1}{h};\ \dfrac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}{F\left(\dfrac{1}{h}+\dfrac{1}{\beta h},\ \dfrac{1}{h};\ \dfrac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}$$

であり上式の第二項の分子を $x\dfrac{dF}{dx}$ と略記すれば

(2.2.46)
$$\begin{aligned}x\frac{dF}{dx}&=\frac{\alpha h\rho^h}{1-\rho^h}\left(1+\frac{1}{\beta}\right)F\left(\frac{1}{h}+\frac{1}{\beta h}+1,\ \frac{1}{h}+1;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)\\&=\frac{\alpha h\rho^h x^{\alpha h}}{1-\rho^h}\left\{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h};\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)\right.\\&\qquad\left.+\frac{1}{\beta}F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+1;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)\right\}\end{aligned}$$



第 2.2 図　アモローゾ分布の回帰関数

と表わせるから結局

$$\eta(x)=\frac{\alpha h\rho^h x^{\alpha h}}{1-\rho^h}\left\{\left(1+\frac{1}{\beta}\right)\frac{F\left(\frac{1}{h}+\frac{1}{\beta h}+1,\ \frac{1}{h}+1\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}-1\right\} \tag{2.2.47}$$

$$=\frac{\alpha h\rho^h x^{\alpha h}}{\beta(1-\rho^h)}\frac{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+1\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}$$

$$\eta^{(\prime)}(x)=\alpha h x^{\alpha h}\left[1+\frac{\rho^h}{\beta(1-\rho^h)}\left\{\frac{\beta+1}{h+1}\frac{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+2\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+1\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}\right.\right. \tag{2.2.48}$$

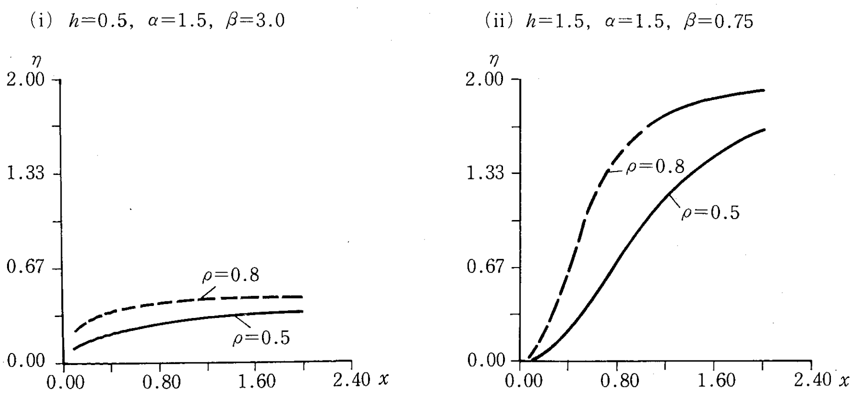
$$\left.\left.-\frac{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+1\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}\right\}\right]$$

となる．故に又

$$\eta(x)+\eta^{(\prime)}(x)=\alpha h x^{\alpha h}\left[1+\frac{\rho^h}{\beta(1-\rho^h)}\left\{\frac{\beta+1}{h+1}\frac{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+2\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}{F\left(\frac{1}{h}+\frac{1}{\beta h},\ \frac{1}{h}+1\ ;\ \frac{\rho^h x^{\alpha h}}{1-\rho^h}\right)}\right\}\right] \tag{2.2.49}$$

が得られる．（第 2.3 図参照）

以上の結果から，$\eta(x)$ の性格を事前に又は標本をもとに推定して，μ_y についての推定精度をたかめることは，一般に定弾力性構造の集団より遙かに複雑になると予想される．然し乍ら変動弾力性構造の集団であっても，適当な層別を行うことによって，各層内を近似的に定弾力的とすることも可能であろう．又それは，パレート分布による何等かの近似，分解，合成，混合或は極限等として表現され得る場合も考えられる．もしそのような関係が見出されるならば，本節及び第 3, 4 節で示されるパレート母集団に対する有意選出法の解析的結果を活用する途が拓かれるであろう．それによって有意選出法は多様な方式を獲得し，その発展の方向を見出すことになるであろう．



第 2.3 図　回帰関数の弾力性係数

2.3. 推定方式の改良とロバストネス

母平均に対する有意選出法に基づく第2推定方式 $\bar{Y}_2$ は，第1推定方式 $\bar{Y}_1$ に $e_{\bar{Y}_1}$ をその補正項として加えることによって得られる．すなわち

$$\bar{Y}_2 = \bar{Y}_1 + e_{\bar{Y}_1} \tag{2.3.1}$$

である．この形式は既に前節で述べたように一般的には $\hat{y}_1, \cdots, \hat{y}_k$ の線形推定量ではない．

我々はその一般論を展開することに先立って，パレート分布を母集団分布と想定して理論的実験を試みることにしよう．以下に於いては母集団をマルディアのパレート第II型分布

$$f(x, y : a, b, p, q, \alpha) = \frac{pq}{(1-\alpha)xy}\left\{\left(\frac{a}{x}\right)^p\left(\frac{b}{y}\right)^q\right\}^{\frac{1}{1-\alpha}} I_0\left[\frac{2\left\{\alpha pq \log\left(\frac{x}{a}\right)\log\left(\frac{y}{b}\right)\right\}^{\frac{1}{2}}}{(1-\alpha)}\right] \tag{2.3.2}$$

$$x > a > 0,\ y > b > 0\ ;\ p > 1,\ q > 1,\ 0 < \alpha < 1$$

$$= 0$$

$$x \leqq a \quad \text{又は} \quad y \leqq b\ ;\ p > 1,\ q > 1,\ 0 < \alpha < 1$$

に従うものとして，第1節でのべた方式でコントロール x を用いて有意選出し，y の平均 μ_y を $\bar{Y}_2$ によって規定することを考える．この場合母平均 μ_y は

$$\mu_y = \frac{bq}{q-1} \tag{2.3.3}$$

と与えられる．[註8)]

今 (2.3.2) 式に於いて α は xy 間の対数モメント相関係数であり，y の x に関する回帰関数 $R(x)$ は

$$R(x) = E(y \mid x) = \frac{qb}{q-1+\alpha}\left(\frac{x}{a}\right)^\eta \tag{2.3.4}$$

ここで

$$\eta = \frac{p\alpha}{q-1+\alpha} < p \tag{2.3.5}$$

である事が知られているから，その近似的回帰関数 $\tilde{R}(x)$ を

$$\tilde{R}(x) = \tilde{\gamma}x^{\tilde{\eta}}\ ;\ \tilde{\gamma} > 0,\ \tilde{\eta} < p \tag{2.3.6}$$

とすることが出来る．このとき

$$\begin{aligned} \mu_{\tilde{R}} &= \int_a^\infty \tilde{R}(x)f(x)dx \\ &= \tilde{\gamma}\int_a^\infty \frac{pa^p}{x^{p-\tilde{\eta}+1}}dx \\ &= \frac{\tilde{\gamma}pa^{\tilde{\eta}}}{p-\tilde{\eta}} < \infty \end{aligned} \tag{2.3.7}$$

となるから (2.1.3) 式で定義される $V(x)$, $W(x)$ は存在し，それぞれ

$$V(x) = 1 - \left(\frac{x}{a}\right)^{\tilde{\eta}-p},\quad W(x) = 1 - \left(\frac{x}{a}\right)^{-\frac{p(q-1)}{q-1+\alpha}} \tag{2.3.8}$$

となる．これにより又 (2.1.4) 式で定義される $\phi(V)$ は

$$W = \phi(V) = 1 - (1-V)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}} \tag{2.3.9}$$

となる．従って容易に

(2.3.10) $$\phi'(V)=\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}(1-V)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-1}$$

(2.3.11) $$\phi''(V)=-\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-1\right\}(1-V)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-2}$$

(2.3.12) $$\phi'''(V)=\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-1\right\}\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-2\right\}\times(1-V)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-3}$$

が得られる．

以上の結果から (2.1.24), (2.1.26) を考察すると有意選出法に基づく第1推定量 $\bar{Y}_1$ は

(2.3.13) $$\bar{Y}_1=\frac{1}{k}\frac{bpq}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}\sum_{h=1}^{k}(1-v_h)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-1}$$

となる．これが (2.1.26) 式と一致することは容易に示すことが出来る．又 $e_{\bar{Y}_1}$ は同様にして

(2.3.14) $$\begin{aligned}e_{\bar{Y}_1}&=\frac{\mu_y}{4\cdot3!k^3}\sum_{h=1}^{k}\phi'''(v_k)\\&=\frac{1}{4\cdot3!k^3}\frac{bq}{q-1}\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-1\right\}\\&\quad\times\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-2\right\}\sum_{h=1}^{k}(1-v_k)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-3}\end{aligned}$$

となる．

第2推定方式は勿論 (2.3.13) 及び (2.3.14) 式の右辺の和で与えられる．

第2.4図は (2.3.2) 式に於いて $b=1$, 又対数相関係数 α が 0.5 と 0.9 の場合について示した．又抽出クラスは $k=10$ に限定した．その結果図のように第2推定方式は実線の示すように

(2.3.15) $$\eta\leqq\tilde{\eta}<p$$

の区間で殆ど完全に近い推定を示すのである．このことは，$\tilde{\eta}$ が必ずしも η をよく近似する必要はなく単に補助変量 x の属性であるパレート係数 p と η の間にとればよいことを示している．つまり (2.3.6) 式による $\tilde{R}(x)$ に比例した有意選出法に関して第2推定方式が理想的な頑健性をもつことを意味する．しかもこの推定方式は p や α の大小にはあまり影響されず，又その変化にも影響されない．但し p が1に近くなると η と p の値が接近するから，$\tilde{\eta}$ を1とすべきかが微妙である．

有効性の強い推定量が x の属性のみで決定されているということは極めて重要である．つまりその場合は調査項目の凡てに対して有効でありうるからである．第2.4図をみれば最大有効性はこの場合大した意味をもたないことが理解されるであろう．有意選出法は少くとも頑健性に関しては有望性に富んだ方法なのである．

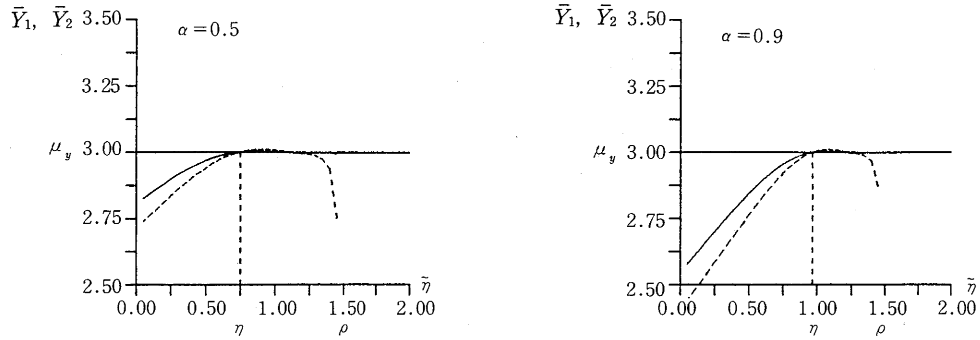
以上のグラフによって得られる結果を解析的に検討してみよう．まず次式に示される λ を導入する．すなわち

(2.3.16) $$\lambda=\frac{p-\tilde{\eta}}{p-\eta},\quad \tilde{\eta}\neq p$$

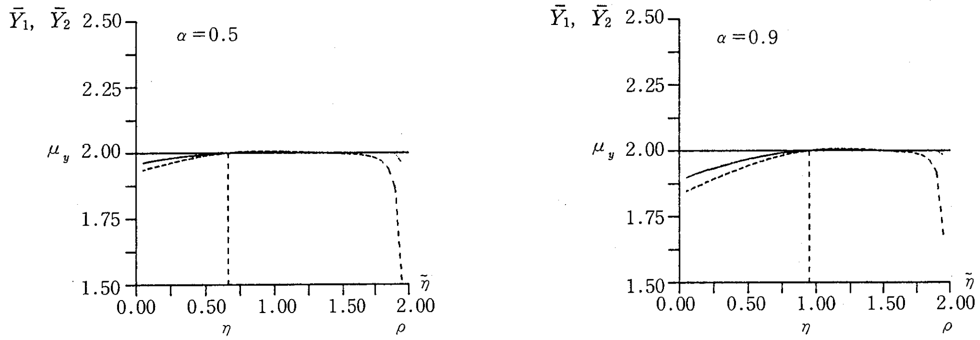
とする．この場合 (2.3.5) 式から容易に

(2.3.17) $$p-\eta=\frac{p(q-1)}{q-1+\alpha}$$

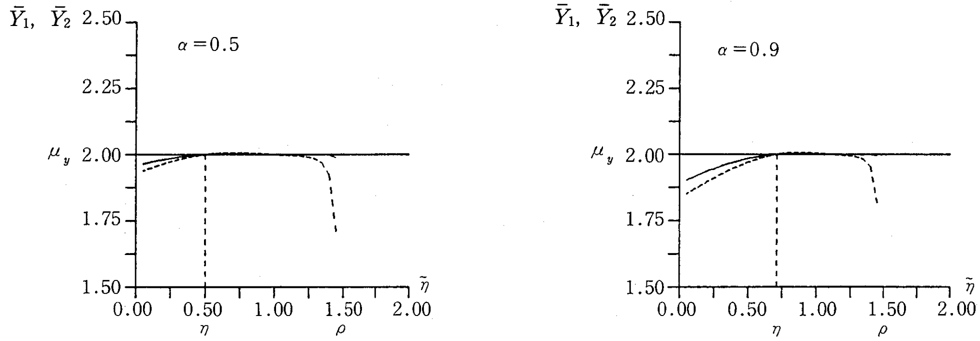
(i) $k=10,\ p=q=1.5$の場合



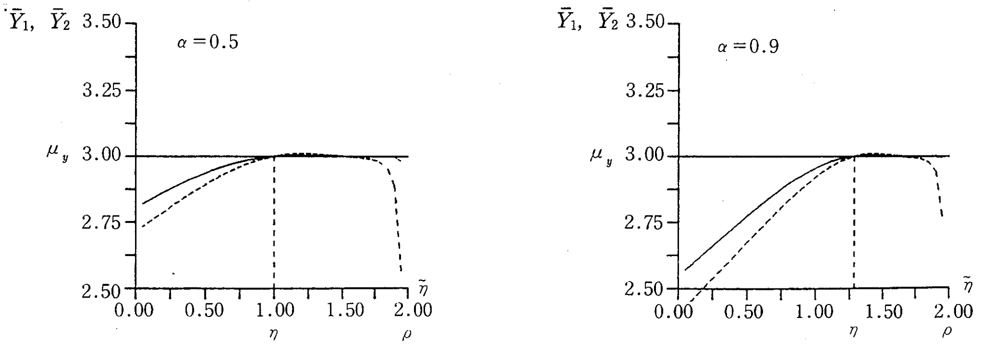
(ii) $k=10,\ p=q=2.0$の場合



(iii) $k=10,\ p=1.5,\ q=2.0$の場合



(iv) $k=10,\ p=2.0,\ q=1.5$の場合



第2.4図 パレート母集団に対する第2推定方式の有効性（実線は第2推定方式 $\bar{Y}_2$ を点線は第1推定方式 $\bar{Y}$ を示す）

が得られるから，(2.3.14) 式は (2.3.3), (2.3.5), (2.3.16), (2.3.17) の諸式を用いて容易に

(2.3.18) $$e_{\bar{Y}_1}=\frac{\mu_y}{4\cdot 3!\,k^3}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\left(\frac{1}{\lambda}-2\right)\sum_{h=1}^{k}(1-v_h)^{\frac{1}{\lambda}-3}$$

と表わされる．今 V_h, v_h の構成を考えると k が充分大ならば，

(2.3.19) $$0<\lambda<\frac{1}{2} \quad \text{すなわち} \quad \frac{p+\eta}{2}<\tilde{\eta}<p$$

の範囲では

(2.3.20) $$\frac{1}{k}\sum_{h=1}^{k}(1-v_h)^{\frac{1}{\lambda}-3}=\int_0^1(1-V)^{\frac{1}{\lambda}-3}dV+O\left(\frac{1}{k}\right)$$
$$=\frac{1}{\frac{1}{\lambda}-2}+O\left(\frac{1}{k}\right)$$

となるから相対誤差 $\gamma_{\bar{Y}_1}$ は

(2.3.21) $$\gamma_{\bar{Y}_1}=\frac{e_{\bar{Y}}}{\mu_y}=\frac{1}{4\cdot 3!\,k^2}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)+O\left(\frac{1}{k^3}\right)$$

となる．従ってこの場合第１推定方式 $\bar{Y}_1$ は λ が小である程，精度の低いものとなる．それは又やや過少な推定値を与える．

従って此の場合は第２推定方式として μ_y を $\bar{Y}_2$ として (2.3.1) をといた形式

(2.3.22) $$\bar{Y}_2=\bar{Y}_1\Big/\left\{1-\frac{1}{4\cdot 3!\,k^2}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\right\}\fallingdotseq\bar{Y}_1\left\{1+\frac{1}{4\cdot 3!\,k^2}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\right\}$$

を用いなければならない．普通の最後の形式は μ_y を単に $\bar{Y}_1$ で置換えた形式に等しい．他方，λ のその他の範囲すなわち

(2.3.23) $$\lambda<0 \quad \text{及び} \quad \lambda\geqq\frac{1}{2}$$

に対しては

(2.3.24) $$\frac{1}{k^3}\sum_{h=1}^{k}(1-v_h)^{\frac{1}{\lambda}-3}=\frac{1}{k^3}\sum_{i=1}^{k}\left(\frac{2i-1}{2k}\right)^{\frac{1}{\lambda}-3}=O(k^{2-\frac{1}{\lambda}})$$

となる．

従って第１推定方式 $\bar{Y}_1$ は

(2.3.25) $$\lambda<0 \quad \text{及び} \quad \lambda\geqq 1 \quad \text{すなわち} \quad \tilde{\eta}>p \quad \text{及び} \quad \tilde{\eta}<\eta$$

のとき，$e_{\bar{Y}_1}$ の値が極めて大きいから，此の場合も第２推定方式も用いねばならない．一方

(2.3.26) $$\frac{1}{2}\leqq\lambda<1 \quad \text{すなわち} \quad \eta<\tilde{\eta}\leqq\frac{p+\eta}{2}$$

の範囲では図の (iv) $\alpha=0.5$ を除き第 2.5 図に示すような比較的小さな誤差が生ずる。結局 (2.3.26) 以外の範囲では，第２推定方式 $\bar{Y}_2$ を用いなければならない．

この場合第２推定方式は

(2.3.27) $$\bar{Y}_2=\bar{Y}_1\left\{1+\frac{1}{4\cdot 3!\,k^3}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\left(\frac{1}{\lambda}-2\right)\sum_{i=1}^{k}\left(\frac{2i-1}{2k}\right)^{\frac{1}{\lambda}-3}\right\}$$

である．

では (2.3.22) 又は (2.3.27) によった場合の誤差 $e_{\bar{Y}_2}$ はどうなるであろうか．まず (2.3.12) 式から容易に

(2.3.28) $$\phi^{(5)}(V)=\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-1\right\}\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-2\right\}$$
$$\times\left\{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-3\right\}\left\{\frac{p(q-1)}{q-1+p}\frac{1}{p-\tilde{\eta}}-4\right\}(1-V)^{\frac{p(q-1)}{q-1+\alpha}\frac{1}{p-\tilde{\eta}}-5}$$

が得られる．従って $e_{\bar{Y}_2}$ は (2.3.16) 式を用いて

(2.3.29) $$e_{\bar{Y}_2}=\frac{\mu_y}{16\cdot 5!}\sum_{h=1}^{k}\phi^{(5)}(V_h)$$
$$=\frac{\mu_y}{16\cdot 5!k^5}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\left(\frac{1}{\lambda}-2\right)\left(\frac{1}{\lambda}-3\right)\left(\frac{1}{\lambda}-4\right)\sum_{h=1}^{k}(1-v_h)^{\frac{1}{\lambda}-5}$$

と表わされる．ここで (2.3.18) 式に対して行ったような考察を加えると

(2.3.30) $$0<\lambda<\frac{1}{4}\quad \text{すなわち}\quad \frac{3p+\eta}{4}<\tilde{\eta}<p$$

に対しては

(2.3.31) $$\gamma_{\bar{Y}_2}=\frac{e_{\bar{Y}_2}}{\mu_y}=\frac{1}{16\cdot 5!k^4}\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\left(\frac{1}{\lambda}-2\right)\left(\frac{1}{\lambda}-3\right)+O\left(\frac{1}{k^5}\right)$$

となる．従ってこの場合 $\bar{Y}_2$ は λ が 0 に近い程精度が悪い．又一般に過小な推定量を与える．他方

(2.3.32) $$\lambda<0\quad \text{及び}\quad \lambda\geqq\frac{1}{4}$$

の範囲では

(2.3.33) $$\frac{1}{k^5}\sum_{h=1}^{k}(1-v_h)^{\frac{1}{\lambda}-5}=\frac{1}{k^5}\sum_{i=1}^{k}\left(\frac{2i-1}{2k}\right)^{\frac{1}{\lambda}-5}=O(k^{4-\frac{1}{\lambda}})$$

となる．

一般に第2推定方式は (2.3.15) の範囲では第2.4図が示すように可成りの精度が期待出来る．特に $\frac{1}{4}\leqq\lambda<1$ に於いて頑健であるから，$\tilde{\eta}$ を弾性値 η よりパレート係数 p に近くとればよいということを示している．

又それは (2.3.29) 式によって理解されるように

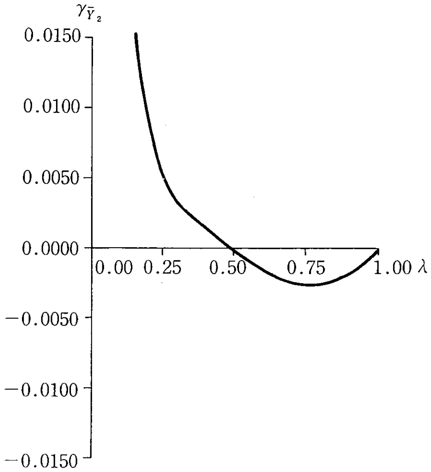
(2.3.34) $$\lambda=\frac{1}{4},\ \frac{1}{3},\ \frac{1}{2},\ 1$$

の各点を境として交互に過小，過大の推定値を与える．第2推定方式は (2.3.22) 及び (2.3.27) 式にみるように λ についての推定を含むものである．それは又その誤差の評価に関してもいえる．

λ の推定が調査後に平均値の第1推定方式と並行して行える事は次節に示される．従って調査企画時に確認されるべき条件は

(2.3.35) $$\eta\leqq\tilde{\eta}$$

である．前節で $\omega=\frac{2}{3}$ とすれば η についての何等の予備知識を必要とせずに (2.2.31) 式が示す



第 2.5 図　第 2 推定方式による誤差の係数（$k=10$）

一定の精度を与える汎用的な $\tilde{\eta}$ が得られることを示したが，そのような $\tilde{\eta}$ が更に (2.3.35) 式の条件を充たすならばそれは極めて推計に適したものといえる．その条件は (2.3.5) 式により

$$\frac{\alpha}{q-1+\alpha} \leqq \frac{2}{3} \tag{2.3.36}$$

が成立することである．

ここでは更に別な実用的方式を示すことにしよう．それは

$$\tilde{\eta}=1 \tag{2.3.37}$$

とすることである．この場合の選出方法は，x に単純に比例して行うことになるから，p についての予備知識も必要とせずに簡易化される．又次章でのべるように経験的資料によって (2.3.37) が (2.3.35) を保障する場合を示すことが出来る．

2.4.　有意標本に基づく統計的推論

第 1 推定方式の誤差 $e_{\bar{Y}_1}$ 又は第 2 推定方式を考察すると，そこに含まれる未知母数は $\eta(x)$ 又は $\delta(x)$ に関係するものである．従って平均値推定の問題は必然的に弾性値の推定を初めとして一般に $R(x)$ の特性量に関する統計的推論の方向に発展する．

そこでまず回帰適合曲線に立戻って，この曲線の基本的特性として適合度 F に着目しよう．F は此の曲線の適合面積を A とすれば $F=2A$ で規定されるものであった．今有意標本に関して次の定義を与えることにしよう．

定義 2.1　「VW 座標平面に於いて原点と k 個の点

$$\left(\frac{h}{k}, \frac{\sum_{j=1}^{h} \phi'(v_j)}{\sum_{j=1}^{k} \phi'(v_j)}\right)=\left(\frac{h}{k}, \frac{\sum_{j=1}^{h} R_j/\tilde{R}_j}{\sum_{j=1}^{k} R_j/\tilde{R}_j}\right);\ h=1,\ 2,\ \cdots,\ k,\ R_j=R(\hat{x}_j),\ \tilde{R}_j=\tilde{R}(\hat{x}_j)$$

を順次に結ぶ折線グラフを $\phi'(V)$ の標本相関曲線とする．この時この折線と均等線 $W=V$ によって囲まれる面積を $\bar{A}_\phi$ とすると

$$\bar{A}_\phi = \frac{\bar{\Delta}_\phi}{4\frac{1}{k}\sum_{j=1}^{k}\phi'(v_j)}; \tag{2.4.1}$$

$$\bar{\Delta}_\phi = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{h=1}^{k}\{\mathrm{sgn}\,(h-i)\}(\phi'(v_i)-\phi'(v_h))$$

である．更に $\bar{F}_\phi$ は $2\bar{A}_\phi$ を与えるものとする」[註9)]

以上の定義によって容易に次の関係が得られる．

定理 2.1 「母回帰適合面積 A は $\bar{A}_\phi$ と次の関係にある．

$$\frac{1}{2}-A = \left(\frac{1}{2}-\bar{A}_\phi\right)\frac{\bar{Y}_1}{\mu_y}+\varepsilon_A+O\left(\frac{1}{k^4}\right) \tag{2.4.2}$$

ここで ε_A は第1節の記号を使うと

$$\begin{aligned}\varepsilon_A &= \frac{1}{4!k^3}\sum_{h=1}^{k}\left\{\frac{\sum_{i=1}^{h-1}\phi'''(v_i)}{k}-2\phi''(v_h)\right\} \\ &= \frac{1}{4!k^3}\frac{\mu_{\tilde{R}}}{\mu_y}\sum_{h=1}^{k}\left[\frac{1}{k}\sum_{i=1}^{h-1}\frac{R(\hat{x}_i)}{\tilde{R}(\hat{x}_i)}\frac{\delta_\eta(\hat{x}_i)}{\hat{x}_i^2 v^2(\hat{x}_i)}\{\delta_\eta(\hat{x}_i)+\delta_\eta^{(\prime)}(\hat{x}_i)\right. \\ &\quad \left. -v^{(\prime)}(\hat{x}_i)-1\}-2\frac{R(\hat{x}_h)}{\tilde{R}(\hat{x}_h)}\frac{\delta_\eta(\hat{x}_h)}{\hat{x}_h v(\hat{x}_h)}\right]\end{aligned} \tag{2.4.3}$$

である．」

証明 第2.1図をもとにして A を表現し，$\phi(V)$ にテイラー展開を適用すると

$$\begin{aligned}A &= \frac{1}{2}-\int_0^1\phi(V)dV \\ &= \frac{1}{2}-\sum_{h=1}^{k}\int_{V_{h-1}}^{V_h}\phi(V)dV \\ &= \frac{1}{2}-\sum_{h=1}^{k}\int_{V_{h-1}}^{V_h}\left\{\phi(v_h)+(V-v_h)\phi'(v_h)+\frac{(V-v_h)^2}{2!}\phi''(v_h)\right. \\ &\quad \left. +\frac{(V-v_h)^3}{3!}\phi'''(v_h)+O\left(\frac{1}{k^4}\right)\right\}dV \\ &= \frac{1}{2}-\sum_{h=1}^{k}\left\{\frac{\phi(v_h)}{k}+\frac{2}{3!(2k)^3}\phi''(v_h)+O\left(\frac{1}{k^4}\right)\right\}\end{aligned} \tag{2.4.4}$$

$\phi(v_h)$ に更にテイラー展開を適用すると

$$\begin{aligned}\phi(v_h) &= \int_0^{v_h}\phi'(V)dV = \sum_{i=1}^{h-1}\int_{V_{i-1}}^{V_i}\phi'(V)dV+\int_{V_{h-1}}^{v_h}\phi'(V)dV \\ &= \sum_{i=1}^{h-1}\int_{V_{i-1}}^{V_i}\left\{\phi'(v_i)+(V-v_i)\phi''(v_i)+\frac{(V-v_i)^2}{2!}\phi'''(v_i)\right. \\ &\quad \left. +\frac{(V-v_i)^3}{3!}\phi^{(4)}(v_i)+O\left(\frac{1}{k^4}\right)\right\}dV \\ &\quad +\int_{V_{h-1}}^{v_h}\left\{\phi'(v_h)+(V-v_h)\phi''(v_h)+\frac{(V-v_h)^2}{2!}\phi'''(v_h)\right. \\ &\quad \left. +\frac{(V-v_h)^3}{3!}\phi^{(4)}(v_h)+O\left(\frac{1}{k^4}\right)\right\}\end{aligned} \tag{2.4.5}$$

$$=\sum_{i=1}^{h-1}\left\{\frac{\phi'(v_i)}{k}+\frac{2}{3!(2k)^3}\phi'''(v_i)+O\left(\frac{1}{k^4}\right)\right\}$$
$$+\frac{\phi'(v_k)}{2k}-\frac{\phi''(v_h)}{2!(2k)^2}+O\left(\frac{1}{k^3}\right)$$

従って A は結局

(2.4.6)
$$A=\frac{1}{2}-\frac{1}{k^2}\sum_{h=1}^{k}\left[\sum_{i=1}^{h-1}\left\{\phi'(v_i)+\frac{\phi'''(v_i)}{4!k^2}+O\left(\frac{1}{k^4}\right)\right\}\right.$$
$$\left.+\frac{\phi'(v_k)}{2}-\frac{\phi''(v_k)}{8k}+O\left(\frac{1}{k^3}\right)\right]$$
$$-\frac{1}{4!k^3}\sum_{h=1}^{k}\phi''(v_k)+O\left(\frac{1}{k^4}\right)$$
$$=\frac{1}{2}-\frac{1}{k^2}\sum_{h=1}^{k}\left\{\sum_{i=1}^{h-1}\phi'(v_i)+\frac{\phi'(v_k)}{2}\right\}$$
$$-\frac{1}{k^3}\sum_{h=1}^{k}\left\{\sum_{i=1}^{h-1}\frac{\phi'''(v_i)}{4!k}-\frac{\phi''(v_h)}{8}+\frac{\phi''(v_h)}{4!}\right\}+O\left(\frac{1}{k^4}\right)$$

ここで

(2.4.7)
$$\bar{A}_\phi=\frac{1}{2}-\frac{1}{k}\sum_{h=1}^{k}\left\{\sum_{i=1}^{h-1}\phi'(v_i)+\frac{\phi'(v_h)}{2}\right\}\Big/\sum_{j=1}^{k}\phi'(v_j)$$

となる事を考慮すると (2.1.24), (2.1.26) 式から (2.4.2), (2.4.3) 式の成立が証明される.（証明了）

従って特に定弾力性係数が存在する場合 $\tilde{\eta}(x)$ も固定して

(2.4.8)
$$\eta(x)=\eta,\ \tilde{\eta}(x)=\tilde{\eta},\ \delta(x)=\delta$$

とすれば次の系が得られる.

系 2.1 「定弾力性構造をもつ集団からの有意選出に於いて $\tilde{\eta}(x)$ を (2.4.8) 式のように，固定すれば，次の関係が成立する.

(2.4.9)
$$\frac{1}{2}-A=\left(\frac{1}{2}-\bar{A}_\phi\right)\frac{\bar{Y}_1}{\mu_y}+\frac{\delta}{4!k^3}\frac{\mu_{\tilde{R}}}{\mu_y}\sum_{h=1}^{k}\left[\frac{1}{k}\sum_{i=1}^{h-1}\frac{R_i}{\tilde{R}_i}\right.$$
$$\left.\times\frac{1}{\hat{x}_i^2 v^2(\hat{x}_i)}\{\delta-v^{(\prime)}(\hat{x}_i)-1\}-2\frac{R_h}{\tilde{R}_h}\frac{1}{\hat{x}v(\hat{x}_h)}\right]$$
$$+O\left(\frac{1}{k^4}\right)」$$

従って δ が充分小となるように $\tilde{\eta}$ を選択するならば，$\bar{Y}_1$ は μ_y のよい推定値となるから A と $\bar{A}_\phi$ 故に F と $\bar{F}_\phi$ は接近する.

更に例題としてパレート分布を取上げてみよう.

例題 1 「目的変量 y と補助変量 x との間にマルディアの第 2 種 2 変量パレート分布の成立を仮定する．この時 (2.3.11), (2.3.12) の各式に (2.3.5) 式の η を適用して

(2.4.10)
$$\phi''(V)=-\frac{p-\eta}{p-\tilde{\eta}}\left\{\frac{p-\eta}{p-\tilde{\eta}}-1\right\}(1-V)^{\frac{p-\eta}{p-\tilde{\eta}}-2}$$
$$\phi'''(V)=\frac{p-\eta}{p-\tilde{\eta}}\left\{\frac{p-\eta}{p-\tilde{\eta}}-1\right\}\left\{\frac{p-\eta}{p-\tilde{\eta}}-2\right\}(1-V)^{\frac{p-\eta}{p-\tilde{\eta}}-3}$$

が得られる．従って更に(2.3.16)のλを用いると

(2.4.11)
$$\phi''(V) = -\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)(1-V)^{\frac{1}{\lambda}-2}$$
$$\phi'''(V) = \frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\left(\frac{1}{\lambda}-2\right)(1-V)^{\frac{1}{\lambda}-3}$$

となる．従って定理2.1によって

(2.4.12)
$$\varepsilon_A = \frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)}{4!k^3}\sum_{h=1}^{k}\left\{\frac{\frac{1}{\lambda}-2}{k}\sum_{i=1}^{h-1}(1-v_i)^{\frac{1}{\lambda}-3}+2(1-v_h)^{\frac{1}{\lambda}-2}\right\}$$

となる．

今λが(2.3.19)の範囲にあるとすれば(2.3.20)を用いて

(2.4.13)
$$\varepsilon_A = \frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)}{4!k^3}\sum_{h=1}^{k}\left[\left(\frac{1}{\lambda}-2\right)\left\{\int_0^{v_{h-1}}(1-V)^{\frac{1}{\lambda}-3}dV+O\left(\frac{1}{k}\right)\right\}+2(1-v_h)^{\frac{1}{\lambda}-2}\right]$$
$$= \frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)}{4!k^3}\sum_{h=1}^{k}\left[\left\{1-(1-v_{h-1})^{\frac{1}{\lambda}-2}+O\left(\frac{1}{k}\right)\right\}+2(1-v_h)^{\frac{1}{\lambda}-2}\right]$$
$$= \frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)}{4!k^3}\sum_{h=1}^{k}\left\{1+(1-v_h)^{\frac{1}{\lambda}-2}+O\left(\frac{1}{k}\right)\right\}$$

更に(2.3.20)の関係を各項に適用すると結局

(2.4.14)
$$\varepsilon_A = -\frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)}{4!k^2}\left\{1+\int_0^1(1-V)^{\frac{1}{\lambda}-2}dV+O\left(\frac{1}{k^2}\right)\right\}$$
$$= -\frac{\frac{1}{\lambda}}{4!k^2}\left\{2+O\left(\frac{1}{k^2}\right)\right\}$$
$$= -\frac{1}{4!k^2}\frac{2}{\lambda}+O\left(\frac{1}{k^4}\right)$$

従って

(2.4.15)
$$\varepsilon_A = O\left(\frac{1}{k^2}\right)$$

である．

他方λが(2.3.23)の範囲にあるとすれば(2.3.24)に準じて

(2.4.16)
$$\varepsilon_A = \frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)}{4!k^3}\sum_{h=1}^{k}\left\{\frac{\frac{1}{\lambda}-2}{k}\sum_{i=k-h}^{k}\left(\frac{2i-1}{2k}\right)^{\frac{1}{\lambda}-3}+2\left(\frac{2(k-h)+1}{2k}\right)^{\frac{1}{\lambda}-2}\right\}$$
$$= \frac{\frac{1}{\lambda}\left(\frac{1}{\lambda}-1\right)\left(\frac{1}{\lambda}-2\right)}{4!k^3}\left\{\sum_{h=1}^{k}\frac{1}{k}\sum_{i=k-h}^{k}\left(\frac{2i-1}{2k}\right)^{\frac{1}{\lambda}-3}+2\sum_{i=1}^{k}\left(\frac{2i-1}{2k}\right)^{\frac{1}{\lambda}-2}\right\}$$
$$= O(k^{1-\frac{1}{\lambda}})$$

となる．

扨，次の定理はパレート母集団に対する有意選出法の有効性を端的に示すものである．

定理 2.2 「例題 1 のパレート母集団からの第 2 章でのべた有意標本による y の母平均の第 1 推性方式 $\bar{Y}_1$ による誤差 $e_{\bar{Y}_1}$ 及び第 2 推定方式 $\bar{Y}_2$ に含まれる λ の 1 次近似 $\bar{\lambda}_1$ は μ_y と $\bar{Y}_1$ が近い時

(2.4.17)
$$\bar{\lambda}_1=\frac{2}{1-\bar{F}_\phi}-1$$

で与えられる．又 $\bar{Y}_2$ が有効である条件は

(2.4.18)
$$-1<F<0$$

である．$\bar{Y}_2$ として更に有効な推定量を得る為には，$\bar{\lambda}_1$ の代りに

(2.4.19)
$$\bar{\lambda}_2=\frac{2Y_2}{(1-\bar{F}_\phi)Y_1}-1$$

を用いて，(2.3.22) 式又 (2.3.27) 式の λ に代入し，それを $\bar{Y}_2$ について解くことによって得られる．又 y の x に関する弾力性係数の標本推定量 $\tilde{\eta}_1$，$\tilde{\eta}_2$ は

(2.4.20)
$$\begin{aligned}\bar{\eta}_1&=p-\frac{p-\tilde{\eta}}{\bar{\lambda}_1}=p-\frac{1-\bar{F}_\phi}{1+\bar{F}_\phi}(p-\tilde{\eta})\\ \bar{\eta}_2&=p-\frac{p-\tilde{\eta}}{\bar{\lambda}_2}=p-\frac{(1-\bar{F}_\phi)Y_1}{2Y_2-(1-\bar{F}_\phi)Y_1}(p-\tilde{\eta})\end{aligned}$$

である．以上における諸精度は例題 1 の ε_A によって決定される．」

証明 例題 1 のパレート母集団に於いては (2.3.9) 及び (2.3.16), (2.3.17) により

(2.4.21)
$$W=1-(1-V)^{\frac{1}{\lambda}}$$

であるから容易に

(2.4.22)
$$F=1-2\int_0^1 WdV=\frac{\lambda-1}{\lambda+1}$$

が得られる．定理 2.1 により $2\varepsilon_A$ の精度で $1-F$ は $(1-\bar{F}_\phi)\dfrac{\bar{Y}_1}{\mu_y}$ で推定されるから

(2.4.23)
$$\frac{\lambda-1}{\lambda+1}=(1-\bar{F}_\phi)\frac{\bar{Y}_1}{\mu}$$

となり，これにより (2.4.17) 式及び (2.4.19) 式が得られる．又 (2.3.15), (2.3.16) を考慮すれば容易に (2.4.18) 及び (2.4.10) が得られる．（証明了）

定理 2.3 で与えられた $\bar{\lambda}_1$, $\bar{\lambda}_2$ を第 1 及び第 2 推定量 $\bar{Y}_1$, $\bar{Y}_2$ 及びその誤差 $e_{\bar{Y}_1}$, $e_{\bar{Y}_2}$ に含まれる λ の代わりに用いることによってパレート母集団の平均値の有意標本による推定方式は完成される．又選出方法の適否は既存データにより (2.4.9) を確認すればよい．F や η も定理2.1～2 により推定されるからパレート母集団を中心とした有意標本による統計推論の形成はその骨格を得たものといえよう．残された課題は寧ろ補助変量 x の解析であり，それは補論に於いて集団構造の解析として取扱われる．

更に次の定理はジブラ母集団に対する有意標本の有効性を示すものである．

定理 2.3 「(2.2.39) 式に示すジブラ母集団の第 1 推定量の誤差に含まれる δ は次式 $\bar{\delta}_1$, $\bar{\delta}_2$ により推定される．

$$\bar{\delta}_1=\frac{\sqrt{2}}{\sigma_x}\Phi^{-1}\left(\frac{1+\bar{F}_\phi}{2}\right),\ \bar{\delta}_2=\frac{\sqrt{2}}{\sigma_x}\Phi^{-1}\left\{1-\frac{Y_2}{2Y_1}(1-\bar{F}_\phi)\right\} \tag{2.4.24}$$

従って又弾力性係数 η は次式の $\bar{\eta}_1$, $\bar{\eta}_2$ により推定される.

$$\bar{\eta}_1=\tilde{\eta}+\frac{\sqrt{2}}{\sigma_x}\Phi^{-1}\left(\frac{1+\bar{F}_\phi}{2}\right),\ \bar{\eta}_2=\tilde{\eta}+\frac{\sqrt{2}}{\sigma_x}\Phi^{-1}\left\{1-\frac{Y_2}{2Y_1}(1-\bar{F}_\phi)\right\} \tag{2.4.25}$$

以上の推定精度は補助定理2.2の ε_A によって与えられる.」

証明 ジブラ母集団に対しては容易に

$$F=2\Phi\left(\frac{\delta\sigma_x}{\sqrt{2}}\right)-1 \tag{2.4.26}$$

が得られる. 従って (2.4.9) に於て μ_y をそれぞれ $\bar{Y}_1$ および $\bar{Y}_2$ とすれば (2.4.24) が得られる. 更に (2.1.32) を適用すれば (2.4.25) が得られる. (証明了)

定理2.3の $\bar{\delta}$ を (2.2.39) 式に代入すれば, ジブラ母集団の第1推定方式の誤差の推定量 $\bar{e}_{\bar{Y}_1}$ が得られる. それは又第2推定量 $\bar{Y}_2$ の推定量 $\bar{\bar{Y}}_2$

$$\bar{\bar{Y}}_2=\bar{Y}_1+\bar{e}_{\bar{Y}_1} \tag{2.4.27}$$

を与える. 従って, ジブラ母集団に対しても有意標本による推定が可能となる.

以上の二つの定理は定弾力性をもった構造に対する有意選出法の有効性を示すものといえる. 然しそのことは変動弾力性をもつ集団構造に対して有意選出法が無効だという事を意味しない. そのような集団構造の解析の手掛りは, さきに筆者が見出した弾力性ベクトル $\boldsymbol{e}(x)$ が与えてくれる.[註10)] このベクトルは補論に示す相互依存関係の一つの表現として, ある種の構造統計表から分布関数形に依存せずに一定の操作によって導かれる. 本来経済学の領域で見出された弾力性概念はそこでは一定に統計的形式の下で表現されている. $\boldsymbol{e}(x)$ によって予想される $\delta(x)$ の性質は, 新しい分布関数を導くことも可能とする. 然し此等は凡て本稿の主題を離れた将来の課題と考えてよいであろう.

本稿にとって重要な事は任意抽出法にとっては非線形回帰構造は, 従来殆ど不毛な領域に属していたということであり, 統一理論の超母集団概念によっても積極的な対策を得ていないということである. そして第1章の抽出企画にしてもその適用範囲は極めて制限されたものであった. 本章によって, 有意選出法は本来非線形回帰構造に適するものであり, 弾力性概念を媒介として現実に接近する方向を与えるものだといってよいであろう.

註1) 回帰適合曲線は後掲 田口 (1984) で詳論する比率曲線の一般化である. その一般的理論は同書で与えた解析的集団の公理系に遡って展開することが出来る. その意味で此の曲線は, ジーニの統計解析方法の発展上に位置する.

註2) メディアールについては, 後掲 田口 (1984) p. 64 参照.

註3) 次の6. で説明するように $e_{\bar{Y}}$ は任意抽出法でいう「偏り」にあたるものかも知れない. その場合標本数を増すと k も大になるから $\bar{Y}$ は「一致性」をもつといえる.

然し「繰り返し操作可能」を前提としない有意抽出法の立場では, 精度は分割数を増せばよくなるのであり, その速度が測定可能であるというだけで充分である.

註4) 二変量のパレート第I型分布については田口 (1984) 参照.

註5) 二変量のパレート第II型分布及びジブラ分布については田口 (1984) 参照.

註6) H.A. Simon 'On a class of skew distribution functions', Biometrika, Vol. 42, 1955, 411-425.

註7) アモローゾ分布については田口 (1984) 参照. いうまでもないことであるが, この分布は例外的にのみガンマ分布として線形回帰構造を示すもので, 一般的には変動弾力性の回帰構造をもつ.

註 8）　註 5）参照.
註 9）　相関曲線については田口（1984）pp. 123-125 参照.
註 10）田口（1984）pp. 138-145 参照.

3.　有意選出法の一般化理論の意義

3.1.　有意選出法と任意抽出法をめぐる論争の回想

前節に於いて経済統計調査に対して有望と思われる方法として定弾性値 η の存在を予想し，それに対する近似値 $\tilde{\eta}$ を求めて (2.2.19) による選出法によって生ずる誤差を検討した．この場合に

$$\tilde{\eta} \longrightarrow 0 \tag{3.1}$$

とすればそれは x の順位による等間隔系統選出法に接近する．序論でのべたジーニ及びガルヴァーニ（1929）の有意選出法は，起点の選出が恣意的であるがほぼ (3.1) の極限状態にあたるものと考えてよい．従ってジーニ及びガルヴァーニは標本選出の段階では何等回帰関係の成立を予想し，又は適用を予定していなかったといってよい．然し彼等は母平均の推定に当っては同時線形回帰モデルを想定しているのである．この矛盾がいわゆる「ジーニの失敗」（伊藤（1926））を招いたといえる．もし選出に際してもその同じ同時線形回帰式を用いて

$$\tilde{\eta} = 1 \tag{3.2}$$

として x の規模に比例した選出を行ったならば，多分結果はもっとよくなったであろうし，第 2 推定方式を用いたならば更に精度は向上したであろう．勿論，ジーニ自身が結論しているように「代表法は相対的概念である．」[註 1] から凡ての特性については期待出来ないかもしれない．ジーニが報告した 1929 年は将に大恐慌の時期であり，イタリアの内外をめぐるその後の政治・経済情勢の変化は彼等にそれ以上の方法論研究を許さなかったかもしれない．この選出法は他方に於いてネイマンが唯一の方法と主張した層化任意抽出法ではない．[註 2] 何故ならば，ネイマンによるとコントロール値を等しくする個体は層を形成するが，その場合この方法は層の一部を選出するものであり，更に選出された層内の個体は全数標本とするからである．これは各層から標本をとるという層化任意抽出法の一般的規定に反する．[註 3] 然し，この選出法は有意選出法の規定に抵触しないし，(3.1) の条件の下でジーニの方法を含むと考えてよい．例えば企業抽出に際して資本金額或いは従業員数が補助変量としてよく用いられるが，此等の値を等しくする個体はその有効桁数のとり方にもよるが，産業別にみると多くは数社に限られる．資本金額 4 億円以上の企業を 10 クラスに分けて選出した時に資本金の上位二桁を同じくする企業数は凡て 10 社以下であり 1 及至 3 社が普通であり，又上位の大半のクラスは 1 社であった．更に最上位のクラスは 1 社が 2 又は 3 クラスの標本となりうるのでそのような場合は，全数調査部分としてそれを別枠とし，その他のクラスを再構成することが出来る．それによって標本数は，概ね当初のクラスの数 k を大巾に上回るものとはならない．各クラス内でその資本金の和のシェアがクラス全体の 1 割を占める程度に隣接した企業を加えても標本数は下位に於いてたかだか 2 倍程度となるにすぎない．これは将にネイマンが与えた任意抽出法が実施出来ない場合にあたる．[註 4] 勿論下に 0 が長く続く歯切れのよい資本金額については慣習上若干多くの企業が数えられるといった例外的な事情は存在する．此等のデータ解析の結果は本稿の続篇で示すが，何れにしてもネイマンが層内抽出を必要とする程ではなく寧ろ全数標本とする方が前提とした回帰関数の構成からみて妥当性をもっている．然もこの場合の選出対象はネイマンの好む個体であり，調査区と違って勝手に組替えを許すような性質のものではない．

結局ジーニ及びガルヴァーニの有意選出法は現実性があり若干の欠陥はあるが改善し得るものであり，ネイマンの批判にこそ独断があったとさえいえるのである．勿論このことはジーニ，ガルヴァーニに疑問がないというのではない．有意選出法の解析に分散や最小二乗回帰方式を持込むことは，当時それに代るべき手法をもたなかったとはいえ，そもそも混乱に導くのである．又例えば，彼の論文（Gini, Galvani (1929)）の pp. 30-32 にまたがる脚註（1）にのべられる分散の計算や分散についての彼等独自の「対称性」の見解は，理解し難いものがある．又 p. 31 の本文に示されるような多項式近似的な非線形関係の例にしても弾力性概念に至らない未熟さが感じとれる．然し彼等は自ら線形モデルや比率の安定性に関する仮定をあくまで彼等の実験的な選出法によって得られたデータを解析する為の手掛りとして用いることを自覚しており，従ってそれが彼等の基本的な結論を誤った方向に導くものではなかった．彼等は又任意抽出法であれ有意選出法であれ限られた標本数の中で凡ての母数について代表的な標本を得ることは不可能なこと，例え特性を母平均の代表に限定しても，それが凡ての種類に亙ることは出来ないだろうと結論した．[註5] 勿論第2章で示したように理論的にはパレート母集団を前提とした時，絶対的代表標本に近いものが得られるが，それは必ずしも現実的とはいえない．

かくして彼等は或意味で正当な結論に達したのであるが，弾力性概念に至らなかった為に，それ以上に具体的な結論を示すことが出来なかった．然し彼等の結論は第1章，第2章の分析を経た今日の我々からみると，弾性値による定量的分類の観点に導くことを暗示しているように思われる．それは外面的にはつまり選出対象となる母集団については第1.1～3表に示したような弾性値による産業分類の方向を示唆するものであり，又内面的にはつまり調査特性については第1.3～4図のような弾性値による項目分類を必要なものとするのである．

このような弾性値解析によって定量的な分類が行われ，それが定性的にみても根拠がある場合は，ある産業群のある項目類に対して代表的な標本を選出することが可成り一般的に可能となる．更に第1.5図のような同一分類項目内にあっても $\delta_\eta(x)$ や $\delta_\eta^{(\prime)}(x)$ を通して階層によって弾性値が著しく変化することが認められ，それが大企業と中小企業或は多重構造による差を反映していると思われる時は，弾力性係数は階層の分類に対しても一定の役割を果たすものとなり，階層別代表標本を選出する基準とすることが出来る．その際，全体的弾力性係数と，それに基づいて構成される回帰適合曲線の解析を通して，対象に則した階層基準を具体的に決定することが可能である．つまり全体的弾力性係数を定弾力性係数としてもつような近似的回帰曲線 $\tilde{R}(x)$ を用いて第2.1図のような回帰適合曲線を構成した場合，この曲線が単一の変曲点をもつならば，第2章第1節の3でのべた理由により，この点に対応する規模 x をもって，大小の規模階層の分点とすればよい．

扨，以上は将に相対的代表性を前提とした階層構成例であるが，その各階層は更に，一定の変動巾の下で特定の弾性値を層の代表値とすることが出来る．

結論として，我々はこれまで述べてきたところにより，一連の解析手続きや調査設計，誤差評価を含む集計の全般に亙って，弾力性概念がいかに統計解析上基本的な役割を果たすものであるかを知ることが出来た．弾力性統計学の定立は，いまや不可避のようにみえる．それは弾力性経済学や規模の経済学等に隣接し，密接に関連することによって，そこに提起される様々な計量モデルを検討し，実証する手段を講ずることが出来よう．

ところで絶対的代表性から相対的代表性への移行は，前章第1節にのべたように回帰適合曲線に関するテイラー展開第1項として導かれた第1推定方式，すなわち (2.1.26) 式で示される比推定形式を前提とする限り必然的と思われるが，これは絶対的代表性概念の存続を全く否定し去ることではない．パレート母集団の場合，第2推定方式をとることによって絶対的代表標本に近いものが得られることは既にのべたが，ここではそれを除外することにしよう．その場

合絶対的代表性の追求は代表法を再び典型調査に転化させる可能性をもっている．次節でそれを解明することにする．

3.2. 有意選出法による接近と典型調査の可能性

既に第 1 章の後半で述べたように不等確率抽出法は，調査項目が多数にわたる時，一様最少分散をみたすものは Godambe の示すように有効サンプルが一致するか又は互いに素なものであった．一致の条件は有意標本を意味し，素の条件は序論にのべたキェールやジェンセンの規定に照して典型標本を意味すると考えられる．[註6)]

有意選出法に於いても，凡ての調査項目についての誤差を適度に抑制する効果的な形式は，第 1 次推定方式の下では一般的には困難であることは，既に前節に於いてジーニの相対的代表性に関連してのべたところである．実際に前章第 2 節〈3〉でのべたように各項目毎に x に関する定弾性値 η_i；$i=1, 2, \cdots, m$ が予想される場合であっても η_i は一般的に互に相異なることがデータ解析によって示されるのであり，その結果 η_i のうちで選出法に用いた近似的回帰関数のもつ定弾性値 $\tilde{\eta}$ との差 δ_{η_i} が大きいもの程一般的には誤差を増大させるからである．然し一方で此等の弾性値は第 1.4 図で示されるように，概ね $\eta=1$ を中心として微少変動を示す（決して $\eta=0$ 近傍ではないことに注意）ことが多く，この点に注目すれば $\tilde{R}(x)$ を単に x に比例する回帰直線とすることによって概してよい精度が保障されるのである．そのことは規模比例確率抽出法が，経済量の推定に於いて屢々層化任意抽出法よりもよい結果を与えるという事実にも符合するのである．

然し前章第 2 節〈4〉でのべた変動する弾力性係数をもつ一般の場合は，必然的に全体的弾性値と局所的弾性値との関係の分析に発展する．もし $e_{\tilde{r}}$ を構成する $\delta_{\eta}(x)$ や $\delta_{\eta}^{(')}(x)$ がある規則性を示す場合は，近似的回帰関数のモデル・チェンジが求められるし，逆に不規則性が認められる場合は全体的弾性値を代表値とするランダムな変量として捉える必要が生ずるかもしれない．又複数の調査項目についての精度の一様性の追求は，更に各項目内及び項目間の全体及び局所弾力性の関係の分析に拡大する必要を与えよう．その結果これまでの回帰モデルや構造モデルを大きく変えることになるかも知れない．其等についてのデータ解析は又本稿の続篇に予定する課題である．

扨，これまで述べたところにより有意選出法の立場に於いても，全調査特性についてよい精度を保障することは一般的には困難であると考えてよい．然し凡ての調査特性の補助変量に関する母集団における全体的弾性値が，互に相異なるものであっても，その各々が母集団のごく限られた一部分における局所弾性値に一致性をみることは充分あり得ることである．例え局所弾性値が変動する性質のものであっても，又それが偶然変動を含むものであってもそのことはあり得ることである．その場合母集団のこの限られた部分を典型標本とし，全数調査することが出来る．これが有意選出法の立場からする典型調査の規定であり，キェールやジェンセンのそれに矛盾しない．又この規定は勿論統一理論よりも具体的である．

キェールは序論で示したように本来典型調査のもつ局所性を嫌って代表法を提唱したのであるが，多数項目に亙る同時精密な調査という課題は，再び典型調査法を浮上させることになるのである．然しここで示した典型調査はかつてのように主観的要素の強いものではない．それは全体的弾性値をある部分における局所的弾性値がよく実現しているか否かを充分に検討し，その保障の下で実施すべきことを要請しているのである．ともあれ我々は歴史の回帰性が，単に代表法の内容に於いてのみならず，典型調査にまで及ぶように思われてならない．しかもこれまでに扱った調査特性とはそれが多項目に亙るものであったとしても専ら母平均に関する推定に限定しているのであり，各項目間の相互関係を示すような複雑な特性については未だに取

り上げる段階にないのである.

ただある程度予想されることは，任意抽出法がモメント統計量，特に分散を中心としてそれなりの標本論を展開することが出来たのに対して，前章にのべた有意選出法は，今後集中統計量[註7)]特に弾力性係数を中心として自己の標本論を展開するだろうということである．それは前者が正規分布を基本として自己の小標本論を形成したのに対して，後者はパレート分布を基本としてそれを達成するだろうという予想に連なる.

3.3. 世帯・企業に対する有意選出法の構想

任意抽出法の一般化としての統一理論が抽象的・形式的であって具体的な集団を扱う統計実務に対して指導性に欠けることは既に第1,2章で指摘したが前章に示した有意抽出法の一般化理論は，経験的事実に基づいて対象に則した具体的な選出方法を呈示することが出来る．我々はそのことを前章第3節の結果をもとにして，まず世帯選出の場合について考察してみよう.

〈1〉 世帯選出についての適用

世帯を標本として選出し，家計簿調査を実施，その支出項目別の総計又は平均を推定する場合を想定する．統計的家計分析に於いては，エンゲル曲線及びエンゲル弾力性の名でよく知られているように，費目別消費支出 y_i；$i=1, 2, \cdots, m$ の消費支出総額 y に関する回帰曲線とその全体的弾力性係数 $\eta_{i\cdot y}$ が対象地域の世帯集団の基礎的特性として，推定の対象とされる.

つまりエンゲル弾力性係数は一般的に

（i） $\eta<0$ のとき劣等品

（ii） $0\leqq\eta<1$ のとき必需品

（iii） $1\leqq\eta$ のとき贅沢品

といった具体的対応を示すことが認められているからである.[註8)] 実際に我国に於いても，総務庁統計局の家計調査資料によって算定すると，第3.1表のような弾力性係数を得ることが出来るのである.

然し実際に世帯を抽出する際は所得を標識として階層を構成して抽出を行うのが通常であ

第3.1表 支出に関するエンゲル弾力性係数[註9)]

計算種類＼支出項目	食料	住居	光熱・水道	家具・家事用品	被服・履き物	保健医療	交通通信	教育	教養・娯楽	その他の消費支出
最小二乗法による推定	0.5688	0.0308	0.5945	1.1799	1.4996	0.6588	0.9554	1.4901	1.2043	1.4853
集中回帰法による推定	0.5764	0.0129	0.5815	1.1742	1.4869	0.6554	0.9419	1.5718	1.2129	1.5091

第3.2表 1か月平均年間収入に関するエンゲル弾力性係数[註10)]

計算種類＼支出項目	消費支出	食料	住居	光熱・水道	家具・家事用品	被服・履き物	保健医療	交通通信	教育	教養・娯楽	その他の消費支出
最小二乗法による推定	0.5685	0.3256	0.0193	0.3425	0.6642	0.8434	0.3751	0.5456	0.8313	0.6840	0.8328
集中回帰法による推定	0.5773	0.3341	0.0075	0.3368	0.6760	0.8539	0.3791	0.5447	0.9012	0.6991	0.8660
$\bar{\eta}$ による推定（局所弾力性係数の加重平均）	0.5615	0.3255	0.0089	0.3075	0.6608	0.8538	0.3666	0.5490	1.0078	0.7105	0.8816

る．従って有意抽出法を前提とする場合に於いても，その選出法の確定に必要なのは，寧ろ項目別消費支出 y_i の所得 x に関する弾性値である．その結果は次の第 3.2 表に示すように一般的に

$$\eta_{i\cdot x}<1 \tag{3.3}$$

が成り立つことを知るのである．

従って前章第 2 節及び第 3 節で解析した処によって，世帯の所得分布がパラメータ (a, p) をもつパレート分布に従い，更に費目別支出 y_i との同時分布が多変量パレート分布に従うものと想定出来る場合は

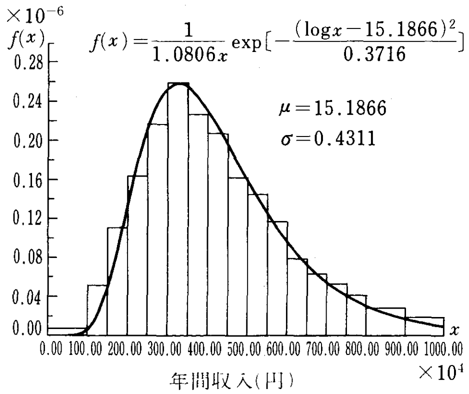
$$\tilde{\eta}=1 \tag{3.4}$$

として選出を行い，(2.3.18) に示す補正項を加えた第 2 推定方式が各 y_i についてよい推定結果を示すであろうことは，第 2.5 図によって伺えるのである．このことは，有意選出法の立場でも所得 x に比例した選出を行い，所得を分母とする比推定形式を第 1 推定方式とすることを意味し，その補正項に独自性が示されることになる．

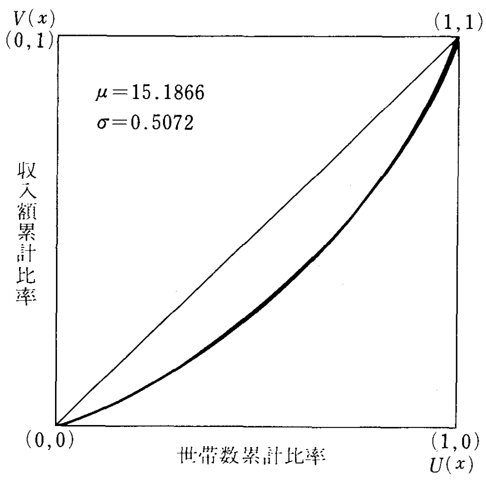
もっとも所得分布をパレート分布がよく適合するという説には異論もあり，その場合は第 3.1 図，第 3.2 図のようにジブラ分布の適合性が高いことから，補正項として (2.2.39) 式で $\tilde{\eta}=1$ とした式を用いることになる．定弾性値をもつという仮定は，第 3.3 図によって或程度理解されよう．[註11] ジブラ分布では又 $\tilde{\eta}=0.5$ も考えられる．又パレート母集団について p についての情報や η についての情報が追加されるならば既に前章で解析したように，高度な推定も行えるのである．勿論，以上をもって統計調査や統計解析の諸問題が尽くされる訳ではない．然し少くとも本稿は任意抽出法を基本とせねばならないという観念に対する一つの問題提起を意味している．

〈2〉　企業選出についての適用

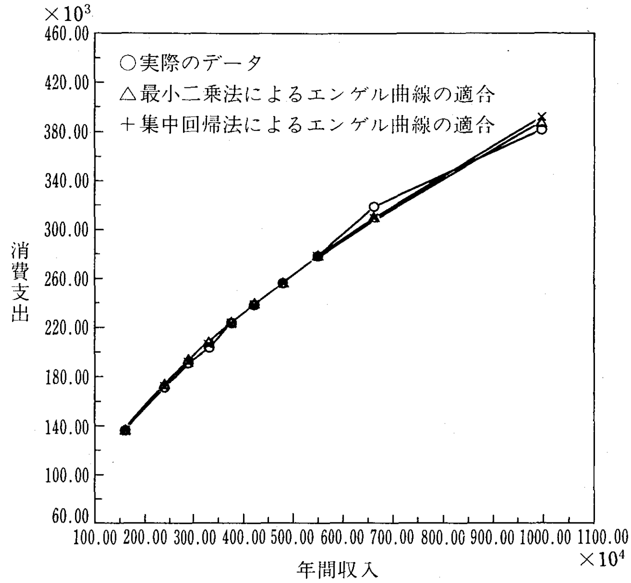
法人企業を標本として選出し，法人企業統計調査にみられるような企業資産，負債及び資本構成等の各構成項目についての総額又は平均を推定する場合を想定する．この場合従来よく用いられる方法は資本金 x を層化基準及び規模比例抽出法における補助変量とすることである．その場合 x と各調査項目 y_i；$i=1, 2, \cdots, m$ との同時分布つまり $(x, y_1, y_2, \cdots, y_m)$ の分布は



第 3.1 図　経験的分布に対するジブラ分布の適合[註11]



第 3.2 図　経験的ローレンツ曲線に対するジブラ分布の適合[註12]



第 3.3 図　年間収入と消費支出の関係[註13)]

〈1〉で取り上げた所得・消費支出分布よりも遙かにパレート多変量分布に近いことが知られている．この場合，各資産・負債等項目の資本金に関する弾力性係数 $\eta_{i\cdot x}$ は各産業につき殆ど凡ての主要項目について 1 より小であり，1 を超える場合も極めて 1 に近い事が認められる．[註14)]

従って此の場合は，前例にもまして資本金に比例した有意選出を行い，資本金を分母とする比推定形式をもとにして補正を加える第 2 推定方式が優れていると思われる．この場合，x として従業員数を用いたとしても適用の条件は満たされている．然しこれ以上の細論は本稿では避けることにしよう．ただ企業選出に際しては，コブ・ダグラス型の生産関数の近似的成立も予想され得るから，今後多変量補助量による有意選出法の適用が有望と思われることを付記しておく．

3.4. 任意抽出法の再吟味とゼロ弾性値

自然科学的事実は，屢々実験により再現される性質をもつものであるから，ランダム・サンプリングの前提となる「繰返し調査の可能性」は抵抗が少いかもしれない．しかしそれとても確認さるべき事実の性格によっては再現可能な期間が限定されるであろう．社会現象については，特に傾向的変化を伴うから，このような「繰返し可能性」は現実的でないとして有意選出法が主張されることが多かった．この従来の論点は決して消滅した訳ではないが，ここでは敢えて取り上げない．

扨，角度を変えてこれまで展開してきた有意選出法の一般化理論の立場では，任意抽出法はどのように位置づけられるであろうか．まず単純任意抽出法によって得られた標本をもとにして母平均を推定する場合を取り上げてみよう．この場合の推定方式は標本平均である．

この方式を第 1 推定方式とする有意選出法は，凡ての x に対して

$$\eta_{\bar{R}}(x) \doteqdot 0 \tag{3.5}$$

となる場合であり，これは等間隔系統選出を意味する．又この場合の推定誤差は (2.1.38) 式に於いて

(3.6) $$\tilde{R}(x)=\mu_{\tilde{R}}=\text{定数}$$

(3.7) $$\delta_{\eta}(x)=\eta(x) \qquad \delta_{\eta}^{(\cdot)}(x)=\eta_{R}^{(\cdot)}(x)$$

(3.8) $$v(x)=f(x) \qquad v^{(\cdot)}(x)=f^{(\cdot)}(x)$$

とおけば得られる．従って

(3.9) $$e_{\tilde{Y}}=\frac{1}{4\cdot 3!k^3}\sum_{h=1}^{k} R(\bar{x}_h)\frac{\eta(\bar{x}_h)}{\bar{x}_h^2 f^2(\bar{x}_h)}\{\eta(\bar{x}_h)+\eta^{(\cdot)}(\bar{x}_h)-f^{(\cdot)}(\bar{x}_h)-1\}$$

となる．この場合第1推定方式が有効である為には，回帰関係の弾性値 $\eta_R(x)$ 自体が常に 0 に近い場合である．すなわち母集団が完全非弾力的状態

(3.10) $$\eta_R(x)=0$$

に近いことが必要である．これは回帰関数が定数となることを意味し，ストカスチックな表現によると無相関の状態にあることを示している．

このような場合凡そ x は補助変量として用いる意味は全くなく従って将に任意抽出法の適用対象といえよう．

第3.3表　売上高の資本金による推定（1981年　資本金4億円以上の企業について）

x^1 による分類

業種	母集団サイズ	標本サイズ	選出比率	変動係数 C	$\tilde{Y}/\mu_y$
建設業	259	36	0.14	1.87	1.08
食料品	204	30	0.15	4.37	1.34
化学工業	392	44	0.11	3.12	0.97
窯業・土石製品	119	21	0.18	2.53	1.12
鉄鋼業	121	19	0.16	0.87	0.92
金属製品製造	110	18	0.16	3.15	0.95
一般機械器具製造	231	34	0.15	4.97	1.01
電気機械器具製造	252	33	0.13	6.10	1.00
輸送用機械器具製造	145	22	0.15	2.44	0.94
その他の製造	110	21	0.19	3.05	0.98
卸売業	465	56	0.12	17.8	1.22
小売業	221	32	0.14	6.01	1.01
不動産業	331	43	0.13	12.2	1.28
陸運業	129	23	0.18	9.55	0.83
水運業	115	23	0.20	4.06	0.96
その他の運輸通信業	149	24	0.16	4.84	0.91
旅館・宿泊	125	23	0.18	4.21	0.92
映画・娯楽	118	21	0.18	6.71	0.80

［註］（$k=10$）

$$\tilde{Y}=\frac{\mu_x}{k}\sum_{h=1}^{k}\frac{\bar{y}_h}{\bar{x}_h}$$

$$\bar{y}=\frac{1}{\sum_{h=1}^{k}N_h}\sum_{h=1}^{k}\sum_{j=1}^{N_h}y_{hj}$$

$$\sigma^2=\sum_{h=1}^{10}\frac{1}{n_h}\left[\left(\frac{1}{N_h}\sum_{j=1}^{N_h}x_{hj}\right)\left(\frac{1}{N_h}\sum_{j=1}^{N_h}\frac{y_{hj}^2}{x_{hj}}\right)-\bar{y}_h^2\right]$$

$$c=\frac{\sigma}{\bar{y}}$$

然し一方では，このような無構造的な集団は果たして解析の対象となりうるのかが問われねばならない．少くとも相互作用や相互依存関係を重要なものとする観点からは例外的な状態といわねばならない．これは将に任意抽出法の立場でパレート分布を例外的としたのと好対照をなすものであり，両者は将に対極的な関係を結ぶことになる．

完全非弾力な状態それは集団認識の未熟な段階や未発達な社会状態に由来するかも知れない．つまりそこでは試行錯誤による反復調査や観察活動が行われると考えられるからである．ストカスチックな統計学は集団論的にみる時このゼロ弾力性をもつ極限の統計学といえる．

然し第1章でのべた不等確率抽出法に関しては一般の有意選出法と同じく何等かの弾力的関係を認めている点で若干事情を異にしている．

この場合両者が同じ弾性値 $\eta(x)$ を利用するならば，一方における $x^{\eta(x)}$ 比例確率抽出法に対して他方における $x^{\eta(x)}$ 比例系統抽出といった差があるが，第1推定方式は同一形式をとる．

然し乍ら，前者は最適推定を目的とする立場で対数回帰残差が x と独立であることを理想とし，その推定方式は不偏であるが，その変動係数は対数相関係数の値により著しく差がある．後者は若干の偏りをもつがロバスト推定を可能とするような補正項とその限界を与えることが出来る．両者の間には異質な概念が含まれており，単純な比較は行い難いが次の諸資料は，何れかの選択を与えるのに有意義な手掛りとなるであろう．そこでは前者の立場での変動係数と，後者の立場での第1推定方式に対する相対誤差が示されている。然しこれは統計利用者としての一つの見方であり，これ以上の推論は避けることにする．

註 1) Gini, Galvani (1929) p. 22 参照．

註 2) Neyman (1934)「以上の理論的考察及びいくつかの例を見て最終的に結論として云えることは，一般的に使える方法としては層化任意抽出法だけが勧められる．」(p. 121 参照)．

註 3) Neyman 上掲論文「唯一つ満たされなければならない条件は，どの層からも最低1個は調査区を抽出することである．」(p. 114 参照)．

註 4) Neyman 上掲論文「…調査区が非常に大きい為に数が少ないときは，ほとんどの二次層には高々1箇の調査区しかないことになる．そうした場合には，もちろん無作為抽出をする余地はない．」(p. 109 参照)．

註 5) Gini 前掲論文第15節「代表性は絶対概念でなく，相対概念である．」(pp. 22-23)にその内容が詳しく説明されている．

註 6) 多賀 (1976) 参照．

註 7) 田口 (1984) 参照．

註 8) N.C. Kakwani, 'Income Inequality and Poverty', Oxford University Press 1980 p. 165 参照．

註 9) 田口 (1984) p. 204 表 7.2 参照．

註 10) 同上表 7.3.

註 11) 田口 (1984) p. 198 図 7.3a)．

註 12) 同上図 7.4a)．

註 13) 同上 p. 199 図 7.5.

註 14) 後掲 田口 (1976) p. 143 第3表及び第4表参照．

補論 集団の構造と相互依存関係[註1)]

センサスに代表される統計調査によって獲得されるクロス・セクション・データは，その多くは第1図に示されるような拡大された構造統計表に構成され利用される．[註2)]

標本調査は通常センサスをもとにして行われるから，基本的にはこの種の結果表の構成を目的とし，更にこのような構造をもつ集団を前提として企画さるべきである．

ではこの種の集団を任意抽出法における集団と明確に区別するものは何であろうか．それは後者が実体分布を伴わない単なる構造統計表を前提とし，又その構成を目的とすることによっ

て生ずる.[註3)]

A.1. 集団構造と相互依存関係の統計的表現

経済学でいう弾力性係数は，相互依存関係の一つの指標とされているが，此の関係はその他にも種々の様相を示すであろうし，又表現出来るであろう.[註4)] 然し統計的にみて相互依存関係は，必ずしも適切な規定を与えられていない.

本節では，相互依存関係に統計的表現を与え，その解析を試みることにする．その為にはまず第1図に示されるような集団構造の規定が必要である.[註5)]

第1図に於いて一般に $\boldsymbol{D}(A_j)=(D_{1j}, D_{2j}, \cdots, D_{nj})$ と書くことが出来る．D_{ij} は A_j 階層の I_i 項目に関する実測値または推定値を示す．分類標識および対象数以外の計量属性欄 $D_i(A_j)$ は実体分布を与える.

以上の記法と第1図を抽象化して解析的集団の公理化を行う.

ある分類標識による階層区分 / 客体の計量属性標識		対象の分類（規模によるカテゴリー）						全　体
		A_1	A_2		A_j		A_m	$A.$
対　象　数		μ_1	μ_2		μ_j		μ_m	$\mu.$
調査項目の標識	I_1	D_{11}	D_{12}	……	D_{1j}	……	D_{1m}	$D_{1\cdot}$
	I_2	D_{21}	D_{22}	……	D_{2j}	……	D_{2m}	$D_{2\cdot}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	I_i	D_{i1}	D_{i1}	……	D_{ij}	……	D_{im}	$D_{i\cdot}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	I_n	D_{n1}	D_{n2}	……	D_{nj}	……	D_{nm}	$D_{n\cdot}$
総　計		$D_{\cdot 1}$	$D_{\cdot 2}$	……	$D_{\cdot j}$	……	$D_{\cdot m}$	$D_{\cdot\cdot}$

第1図　経済統計表の抽象的一表現

定義 1. 「集合 Ω における σ 加法族 $\mathscr{A}$ に属する任意の集合 A に対して，測度 $\mu(A)$ およびベクトル値集合関数 $\boldsymbol{\varDelta}_n(A)$ が与えられるものとする．このとき $\boldsymbol{\varDelta}_n(A)$ の各成分 $\varDelta_i(A)$ が完全加法的であるとき $(\Omega, \mathscr{A}, \mu, \varDelta_n)$ を n 次元の統計解析的集団または，単に n 次元の解析的集団 H_n と定義する.」勿論 $\boldsymbol{\varDelta}$, $\varDelta_i$ はそれぞれ $\boldsymbol{D}$, D_i に対応した抽象的規定である．更に今解析的集団の n 個の調査標識を n 重の分類標識に用いると次の n 次元の集団構造を与えることが出来る.[註6)] つまりそれは各々の実体分布をもとにしてローレンツ曲線（集中曲線）が得られることを想定したものである．つまり

定義 2. 「n 次元の解析的集団において Ω を n 次元ユークリッド空間 R_n とし，また $\mathscr{A}$ をボレル集合族 $\mathscr{B}$（全ての区間から生ずる σ 加法族）とする．更に $\mu(R_n)=1, \varDelta_i(R_n)=1$；$i=1, 2, \cdots, n$ とする．このとき (1) ボレル集合 B（$\mathscr{B}$の要素）に対して，$S_n(B)$ の全体が特異点のない閉曲面の内点および面上の点となり，(2) その曲面は可付番個の点を除き，$I_p=[-\infty, x_p]$ または $I_p^c=(x_p, \infty]$；$p=1, 2, \cdots, n$ の決める区間 B_x に対して $S_n(B_x)$ の全体によって与えられ，(3) かつすべての p に対して 2 次元ベクトル $(\mu(B), \varDelta_p(B))$ が平面上の凸集合をなすとき，この集団を特に n 次元の集中構造 K_n という．またこの場合 $S_n(B)$ を特に集中ベクトルま

たは単に集中量 $C_n(B)$ という。」

その上でローレンツ曲線の一般化としての集中曲面を次のように定義する.

定義 3. 「n 次元の集中構造は, 定義2により集中量 $C_n(B)$ の全体に対して特異点のない境界曲面をもつが, この曲面を n 次元完全集中曲面, またその曲面で囲まれる体積を n 次元集中度という. 更に $C_n(B)$ の全体を含む $n+1$ 次元ユークリッド空間 R_{n+1} 内の領域を集中空間という.

特に $n=1$ のとき $C_n(B)$ の境界曲線を完全集中曲線, またその内部の面積を集中度, 集中曲線を含む領域を集中平面という.」[註7)]

私見によれば, 定義1に於いて Ω を一般に k 次元ユークリッド空間とし, $n=0$ つまり μ のみが与えられた集団（ゼロ次元の集団構造）を対象とした時, ピアソン流の解析が成立するのであり, そこでは専ら分類標識間の関係が解析の対象となる. 従ってその関係は専ら度数分布を介して把握されることになる. 今このようにして把握された関係を一般の相関関係というならば, 相互依存関係とは $n\neq 0$ であって, 調査項目の標識間の関係を意味する. 従ってこの場合の解析は単に度数分布のみならず実体分布 Δ を介して把握されることになる.

一般の相関関係及び相互依存関係の以上の規定に従うと, 調査標識と分類標識が一致する定義2で与えた集中構造に対しては, 同一標識間の関係が異なる観点の下で別様に把握されることになる. そして他方ではパレート分布のように平均値以外の高次モメントを所有せず従ってピアソン解析を許さず, 集中解析の適用のみを認める場合も生ずるのである. これと対置する場合として平均値が0のガウス分布を挙げることが出来よう.

ピアソンに初まるストカスチックな解析とジーニに初まる集中解析は, 以上のべた二つの観点をそれぞれ代表するものであるが, 其等は相互に対応はするが相異なる独自の表現形式を与える. その詳細は田口（1984）に詳しいが, そこでは上に規定した相互依存関係が更に弾力性係数に独自の形式と幾何的性格を与えることが示されている.

A.2.　相互依存関係の諸形態

前節に於いて相互依存関係は専ら定義2に示す集中構造をもとにして考察された. 相互依存関係は $n\neq 0$ である限り, 集中構造に対してのみならず定義1に示した一般の解析的集団に関しても考えられるのである. その場合それは集中構造に対するよりも遙かに多様な形式と豊富な内容を持ち得るのである. 又多種な応用例を与えると予想される. 例えば産業関連表は, 非計量的分類標識としての産業を用いた場合の相互依存関係の一表現と考えてよいであろう. 又寄与度・寄与率は, 時間をはさんでの相互依存関係ともいえる.

本稿で有意選出法の基礎とした回帰適合曲線は又, 相互依存関係の観点に基づいた特殊な応用例とする事が出来る. つまりそれは X を分類標識とした時調査項目 Y に関する理論と実際との相互依存関係の一表現といえるからである. この曲線は又 y と $\tilde{R}(x)$ に関する二次元集中構造における比率曲線という事も出来る.[註8)]

二標本問題に於いては実体分布の累積に替えて, 累積度数分布を用いた同種の曲線表現があるが, これも相互依存関係の特殊例として扱うことが出来る.

然しこれ以上集団構造や相互依存関係についてふれることは既に本稿の予定を越えるものとなるであろう. これについては別に稿を準備せねばならない程広汎である.

従ってここでは単に次のような回帰適合曲線に準じた曲線の構成と解析が, それぞれ分布モ

デルの吟味及び変動弾力性係数モデルの解析の上で有用となることを指摘するにとどめる．すなわち

(1)　X の経験的累積度数分布 $F(x)$ と理論的累積分布モデル $\tilde{F}(x)$ を用いて上述の標本問題の例に準じた分布適合曲線を構成すること

(2)　Y の X に関する経験的累積変動弾力性係数 $\eta(x)$ と理論的累積弾力性関数 $\tilde{\eta}(x)$ を用いて弾力性適合曲線を構成すること

である．

結論として有意選出法は相互依存関係に則した標本調査法なのであり，それによって経済統計に則した方法なのである．更に蛇足乍ら相互依存関係の視点からの不確実性の追求も可能なのである．

註1)　本補論は第2章を補足する目的で，特に回帰適合曲線の発想の源に遡ろうとするものである．その際解説の基礎として集団構造の詳細については田口（1984）を参照されたい．

註2)　工藤弘安「統計学」地方公務員研修選書9（昭和55年，学陽社）によると構造統計とは特定標識についての観察結果を同質的ないくつかのカテゴリーに分類し集計することにより統計集団の内部構造を明らかにするデータである．又構造統計の例として職業別人口，経営耕地規模農家数等があげられている．(p. 40)

註3)　米沢治文「経済統計計量分析」(1972，日本評論社) においては上例のような度数分布のみでは集団の構造分析には不充分であるとする．経済統計の場合は，いまひとつ分配の問題が介在するから実体分布 (p. 60) が必要であるとする．

註4)　島津亮二「経済学辞典」，1979 岩波書店 p. 853「弾力性とは，相互依存関係にある変数間で，ある変数の変化の度合を総体的な比率で示すものであり…」松嶋 (1985) p. 324「彼はいわゆる「因果関係論」を拒否し，相互依存関係の総体をこそ重視していたから．」

註5)　集団構造については註1) を参照されたい．

註6)　三潴信邦・関弥三郎「経済統計論」(昭和60年，有斐閣ブックス) によれば「すべての統計調査はその結果を分類集計することを目的としている．調査表に取り上げられた調査標識(具体的には調査項目) は分類・集計の段階では分類標識として改めて登場する．」(p. 68 分類標識と標準化)

然し勿論凡ての調査標識について実体分布が得られる訳ではない．例えば質的標識区分 (産業・職業等) をもつ場合がそうであり，又計量標識であっても，単なるステイタス・シンボルのようなものは実体分布をもたないであろう．

註7)　以上の諸定義は田口 (1984) 第2章，統計解析の基礎概念 pp. 23-50 による．

註8)　集中構造や比率曲線に関しては田口 (1984) 第4章に詳しい．

む す び

標本調査論は，人口統計や社会統計に於いて重要であるばかりでなく，経済統計の領域に於いても欠くことのできない調査方法を与えるものである．従って此の方法は単に数学的方法論的観点のみからだけでなく，その適用領域の側から，その意義や効果が問われねばならない．

本稿における有意標本論の展開は，単に統計調査の設計に止まらず，推論，解析の各面に及ぶことになるが，その際いたるところで見出されるのは，全域的及び局所的弾力性係数（弾性値）との関係であり，それが問題の解析を可能とするキイとしての役割を果たしている点である．このことは従来の標本調査論に於いては，特に注目されることのなかった事実である．その反面に於いて，この条件が特にミクロ経済分析の領域に於いて親しみのある計量概念であることも周知の事実であろう．弾力性概念は本質的に非線形な集団現象への対応として成立したのでありストカスチックな方法概念を超えた処にある．

統計解析や標本調査論における弾力性係数の役割の発見はミクロ経済学との垣根を一気に乗り越えさせる感がある．これによって弾力性概念は単に経験的分析手法としてのみならず，統

計数学的観点からも注目されてよいであろう．それは現に高階化した形で有意選出法の推定方式や誤差評価を可能にしたのみならず，従来主観的な概念としてのみ考えられて来た典型調査を客観的に規定する一つの手掛りを与えるからである．この弾力性概念は又既に筆者(1976)により指摘されたようにベクトル解析を多用する統計的集中解析とも密接な関係をもつものである．本稿を契機として筆者はストカスチックなモメント解析法とは別に敢えて「弾力性解析法」の成立を予想したいのである．それはジーニを初めとするかつての大陸派の統計方法論に解析的支柱を与えるであろう．特に所得分布の解析に弾力性概念の導入が適していることは，既に拙著（1984）で示されている．

弾力性概念と共に本稿で中心的な役割を果たすのはパレート分布である．本稿で一般化された形で提起された有意選出法はパレート分布に於いてよき適用対象を見出すのでありそれは任意抽出法における正規分布に対置出来ることは既にのべた．このような対置は弾力性係数と分散についても見出すことが出来る．かつてアモローゾはジーニの方法を非ユークリッド的と評価したが，もし平行線の公理に中心極限定理をみたてるならば，この評価は理解出来るような気がする．少くともパレート分布を中心として統計推論が築かれるならば，それはノン・ストカスチックな推論といえるであろうことは可成り確実なことである．パレート自身彼の純粋経済理論を力学的と述べていることも一つの根拠となるが，レヴィーを初めとする確率論者の側に於いてもこれを自己本来の解析対象に加える程の自信と勇気があるであろうか．パレート分布に対するストカスチックな解析結果は順位統計量に関するものが多く，そこに一つの独自性が認められる．

ところでパレート分布を中心とする統計推論が成立したとしても，これをもって筆者はパレートの理論が凡て正当であったと主張するものではない．彼のいう「相互依存関係」は統計解析に対してどのような立場を与えるであろうか．非ゼロの弾性値や曲線相関はそのような関係の反映であろうか．筆者は充分な理解をもたない．それと共に非ストカスチックな立場を直ちに決定論とし，「因果関係論」に解消することも出来ない．[註1)] この点ではパレートが単純な決定論を拒否する立場で「相互依存関係」を重視した事に関心がもたれる．「相互依存関係」が「相関関係」とも異なる事は彼の「力学的」且つ「非ストカスチック」な理論構成から推察出来るような気がする．

更にパレート分布を彼の「集中解析」の中心においたジーニの統計方法論は，彼が一方ではパレートを主観的と評し，又非均衡論的な「経済病理学」を展開している点から改めて関心が持たれるのである．彼の提案になる多種多様の相互関係を表わす統計量は，パレートの「相互依存関係」をどのように踏襲，継承或いは変形しているのであろうか．或いはパレートの論理を補完する立場で得られたものがあるであろうか．或いは又それに対して「残基」と「派生」の関係にあるのであろうか．[註2)] 何れにしても単純なストカスチックな立場の「相関関係」と異なって興味のつきないものがある．勿論以上の両者の関係を単に政治的角度からファシズム理論の系列としてのみ捉える立場は論外である．

以上のように有意選出法を基礎とした今後の統計解析は，興味ある多くの課題を与えてくれる．本稿は単に統計調査論上の解析にすぎないが，これによって従来とかく疎遠で相交わることのなかった調査企画とミクロ経済分析がその距離を縮めて，理論と実証の立場で連携し，相互の機能を促進させる一つの緒となるならば筆者の望外の喜びとせねばなるまい．

註1） 松嶋［16］p. 324 参照．

註2） 例えばC. Gini, Statistical Methods, Biblioteca del 'Metron', Istituto di Statistica e ricerca sociale, Corrado Gini, Universita degli Studi di Roma, Roma 1966. に集成されている．

後記並びに謝辞

筆者の標本調査法との出会いは，昭和30年国富調査の準備期に遡る．林知己夫前統計数理研究所長（当時第三研究部長）の鞭撻下のサンプリングの学習は直ちに調査設計・実査に連ながり，息つく間もない強行軍であった．

経済企画庁（当時審議庁）調査部統計課に出向いての国富調査の客体選定作業では，否応なしに地方公共機関・営利・非営利法人，個人企業等の各種の全国的大規模集団の資料と対面せねばならなかった．

この苦役の体験から筆者は本稿の萌芽ともいえる有意選出法の発想を得ることが出来た．それはローレンツ曲線にテイラー展開を適用するという幼稚な試みであったから，今からみると可成りジーニの方法に近いものであったが，当時の浅学非才の力量をもってしては忽ち方向を見失ったのは当然であった．

其後ハンセンの不等確率抽出法や統一理論についての知識を得て，筆者は昭和45年度国富調査に於いては規模比例確率抽出法の適用を提案したが，同意を得るまでに至らなかった．国富調査に続き，法人企業統計調査の標本設計の機会を得た筆者は，本稿の第1章の初めにのべた不等確率抽出法の着想を得て，その一部を実査に移すことが出来た．

昨年度（昭和60年度）日本統計学会で共同テーマとして再び「国富調査の標本設計について」報告の機会を得た筆者は，それを契機に以後懸案の有意選出法の再検討に踏切った．

ジーニ以前の標本調査論は，代表法として任意抽出法と有意選出法を含むものであったが，此の時期に活躍した一方におけるフィッシャーに対するキェール，ジェンセンについての見解は馬場吉行氏（元甲南大学教授）の名著（1969）で克明に研究，紹介，解説されており，それは本稿の序論の骨格をなしている．

我国におけるジーニ及びガルヴァーニの有意選出法（1929）の紹介は，多くはネイマンのそれに対する批判的論文（1934）から取材したものであるが，伊藤教授の論文（1926）が最も詳細なものといえよう．氏の批判力は鋭利でこれは著者にジーニ，ガルヴァーニ及びネイマン論文への関心をつのらせ，再検討への引金となった．

ネイマンの論文の全容を知る上では，統計数理研究所の岸野洋久助手の助力を仰ぐことが出来た．又ジーニ，ガルヴァーニの論文については著者がイタリー派遣研究員としてローマ滞在中に世話になった現ローマ大学統計学部長のリッズィ教授の手を再び煩わしてこれを入手出来た上に一橋大学博士過程を修了し，母校のボッコーニ大学で教鞭をとることが内定したコラード・モルティーニ氏の助力によってその全容を知ることが出来たことは全く幸便という他はない．

此等の論文の検討を終えて筆者は第3章で指摘したようにネイマンの批判は，疑問を含むものであり，又ジーニの論文は直接的には我国で殆ど全く紹介されていない事に気付いた．それは不完全ではあるが含蓄に富むものであり，本稿の骨格をなす第2章は此等の論文の検討期間中に徐々に成熟したといってよい．それによって61年3月に行った「官庁統計における調査・集計・解析に関する研究」（60-共研-23）集会において「標本調査論における二つの方法」と題して中間報告を行うことが出来たのである．

又初稿についてのレフェリー氏は，統一理論に基づいたロイヤルの論文（1970）とモデル・デペンドをめぐるストカスチックな標本論争 Hansen et al. (1983) を指摘されて筆者の興味を唆り，研究を促進させた．そのことは筆者の感謝にたえぬ処であるが，敏感な読者は既に十分に察知されたであろうように，この論争におけるモデルとはストカスチックな性格のものであり，又ロイヤルが統一理論に基づいていうところの有意選出法とは，ランダムの否定としての

形式的消極的な規定であり，決して具体的な適用条件や手続きに基づいた規定ではないといえよう．

因みに統一理論については多賀保志教授のよく整理された労作（1976）によって得る処が大であった．又パレート研究を集大成された松嶋敦茂教授の啓蒙的な好著（1985）に接し，そこではパレートのダイナミックな観点が浮彫りにされており，それがストカスチックな観点とは明かに異質なものであることを知らされるのである．勿論，純粋経済学は純粋自然科学として分類される（松嶋（1985）p. 176）等のパレートにおける論理的観点は多くの問題を含むものであり，それによって直ちに物理的力学系を経済統計に導入するといったことは無意味であるが，ストカスチックな観点をもって統計方法論を一元化しようとする数理統計学者の一部の思考に対して一つの反省素材を与えるものであろう．

然し乍ら筆者の有意選出法の再提案は，決してネイマンが且て反対の立場で行ったように，ランダム・サンプリングを全面的に否定するものではない．コントロールとなるべき有効な事前情報が見出せない場合，又分布がパレート分布のような高度の歪みをもたず，中心極限定理の適用が比較的無理を伴わない場合は，ランダム・サンプリングは勿論極めて有用な方法でありうる．

筆者の結論は結局初期の標本調査論にみられる有意選出法と任意抽出法の並存論に回帰するのであり，そこに追加したものがあるとすれば，有意選出法をやや一般的にフォーミュレイトした事であり，又それぞれの方法の適用条件をやや鮮明にした点にあるといえよう．そしてその場合に於ても筆者は色々な機会に数々の教示を受け励まされたのであり，此の機会に関係諸氏に深く感謝せねばならない．それと共に筆者は本稿の諸計算資料を作成し，原稿の全体について整理された山田晴美嬢に感謝してやまないものがある．

参 考 文 献

馬場吉行（1969）．増補，標本調査法の基本問題，有斐閣，昭和44年，3版．

Bowley, A.L. (1926). Measurement of the precision attained in sampling, *Bull. Inst. Inter. Statist.,* **22** (1).

Fisher, R.A. (1925, 1958). *Statistical Methods for Research Workers.*

Gini, C. and Galvani, L. (1929). Di una applicazione del metodo rappresentativo all' ultimo censimento italiano della popolazione, *Annali di Statistica,* serie VI, **IV**.

Hansen, M.H. and Hurwitz, W.N. (1944). A new sampling of the population: Sampling principles introduced into the Bureau monthly reports on the labor force, *U.S. Bureau of Census.*

Hansen, M.H., Madow, W.G. and Tepping, B.J. (1983). An evaluation of model depend and probability sampling inferences in sample surveys, *J. Amer. Statist. Ass.,* **78**, 776-805.

畑村又一・奥野忠一（1949）．標本調査入門，小石川書房．

Hayashi, C., Maruyama, F. and Ishida, M.D. (1951). On some criteria for stratification, *Ann. Inst. Statist. Math.,* **11**, No. 2, 77-86.

林知己夫・多賀保志・高倉節子（1955）．全数調査と抽出調査を併用する場合のサンプリングについて，統計数理研究所彙報，**2**, 第2号，11-24.

林知己夫・丸山文行（1949）．ある層化法について，統計数理研究所講究録，**4**, 第10号，399-401.

林知己夫（1951）．サンプル調査はどう行うか，東大出版．

伊藤陽一（1926）．有意選出法の検討，北大経済学，第**4**号．

Jensen, A. (1926). The Application of the representative method, *Bull. Inst. Int. Statist.,* **22** (1).

Kiaer, A.N. (1896). Observation et experience concernant des denombrements representatifs, *Bull. Inst. Int. Statist.,* **9** (2).

増山元三郎（1950）．連続分布で近似できる有限母集団で一半は全部，他半は一部調査する場合の境目の推定法，統計数理研究所講究録，**5**, 第2号，85-87.

松嶋敦茂（1985）．経済から社会へ――パレートの生涯と思想――，みすず書房．

Neyman, J. (1934). On the two different aspects of the representative method; The method of stratified sampling and the method of purposive selection, *J. R. Statist. Soc.,* **97**, 558-625.

Royall, R.M. (1970). On finite population sampling theory under certain linear regression models, *Biometrika,* **57**, 377-387.

Taga, Y. (1967). On optimum stratification for the objective variable based on concomitant variables using prior information, *Ann. Inst. Statist. Math.,* **19**, No. 1, 101-130.

Taga, Y. (1971). On the convergence of optimum stratifications for empiric distribution function in univariate case, *Ann. Inst. Statist. Math.,* **23**, No. 3, 355-364.

多賀保志 (1976). サンプリング調査の理論, サイエンス社.

田口時夫 (1975). 日本における法人企業の標本抽出について, 統計数理研究所彙報, **22**, 第2号.

田口時夫 (1976a). 最有効確率抽出法と回帰関数比例確率抽出法, 統計数理研究所彙報, **23**, 第2号.

田口時夫 (1976b). 最有効確率抽出法とその適用について, 数理解析研究所講究録, **272**, 京大数理解析研.

Taguchi, T. (1978). On an unbiased consistent and asymptotically efficient estimation of Gini's concentration coefficient, *Metron,* **XXXVI**, No. 3-4.

田口時夫 (1984). 経済分析と多次元解析, 東洋経済新報社.

Wakimoto, K.; Stratified random sampling,

(I) (1971) Estimation of the population variance, *Ann. Inst. Statist. Math.,* **23**, No. 2, 233-252.

(II) (1971) Estimation of the population covariance, *Ann. Inst. Statist. Math.,* **23**, No. 3, 327-338.

(III) (1971) Estimation of the correlation coefficient, *Ann. Inst. Statist. Math.,* **23**, No. 3, 339-354.

Reexaminations of the Theory of Sample Survey
—— On the foundation of purposive selection ——

Tokio Taguchi

(The Institute of Statistical Mathematics)

Some historical problems in sampling theory are discussed. An extended method of purposive selection is suggested and its theoretical foundation is researched.

The coefficients of elasticity of regression curves on regressor are introduced and discussed to express the estimate error of the purposive selection. The sample selections from the Pareto and Gibrat populations are precisely discussed, too.

The historical papers by C. Gini and J. Neyman on representative methods are reexamined. Some practical forms of purposive selection are exhibited as examples.