

# 競争システムにおける統計

— イタリア学派の展望 —

統計数理研究所 田 口 時 夫

(1968年8月受付)

Statistics for Competitive System  
— On the Italian Statistical School —

Tokio Taguchi

We introduce the methods of Italian statistical school in the first part of this paper. But some one of them, for example, the measures of skewness and sharpness of concentration curve and generalized mean difference, etc.; and their representations were devised by the author.

In the second part we propose to found those methods by the dynamic notions: the powers of competition, oligopoly, monopoly and so on, as the forces mutually working with grade and between grades under the competitive system.

In the third part, we treat the applications to some technical problems on the statistical researches.

The Institute of Statistical Mathematics

## 序 論

C. Gini 教授を追悼する一文で V. Castellano [3] はこの 'vanguard of ideas' の創見になる 'Italian statistical school' の性格を K. Pearson に基礎づけられた 'Anglo-Saxon statistics' との比較に依りて特徴づけている。そこではこの二大潮流はそれぞれ特異な見解と数学的方式を有し乍ら、全体として後者が経験的事実を離れて抽象数学化する傾向を有するのに対し前者が常に経験的事実の消化と吸収とに忠実であることに一つの特徴を見出しているのである。それはいわば 'deductive' に対する 'inductive' な process を辿らせているのである。勿論この主要傾向は後者が統計的観測についての技術的努力を無視したりその種の寄与が皆無であるということの意味する訳でも両者の補完的な側面を無視する訳でもない。寧ろその当否は別として前者における 'operational research' への傾向から両者の対象を通じての接近を近い将来に予見しているのである。

此等の事實は C. Gini 教授<sup>4)</sup>自身の Italian statistics の説明に於てより顕著である。その中に我々は教授の数理的技術的成功や完備な公理化への関心より、対象の本質の科学的洞察への関心の強さを認めない訳にはいかない。

だが今日では我々はこうした認識主体の側における伝統、関心、思想の中に両者の特徴を認めるにとゞまらず、客体の側に即ち両者が接近した客観的对象の中にその必然性と方法の妥当性を、業績評価の基準を、又両者の差異と一致と補完性を見出しうるように思われる。

翻って従来我国に於ては、前者の紹介、評価、研究等が可成り不足しているかのように思われるので、筆者は敢てこゝに報告を試みるのであるが、勿論それによって 'One ends up by not seeing the wood for the trees' を解消したという自負はない。即ち筆者に於て両者の差異が一面に於ては有機的現象に対する無機的現象の差であり Pareto 的歪み型分布関数に対する正規型分布関数の数学的性質上の差であり、より抽象的には 'qualitative' な側面に対する quantitative な側面の差に還元出来るように思う一方でそのあまりに皮相な観察に過ぎないのではないのではないかと危惧を必ずしも免れていないのである。

## I. 基礎概念 [6]

### 1. 分配 (share distribution) 関係と価値法則

一般的に所得、資産等に関する国民経済統計では、しばしば度数或は確率法則の存在よりも社会的分配法則の支配とその形態が直接的重要性を有するであろう。勿論しばしば前者を前提として当該システムに関する適当な条件下で部分的・一時的に後者のよきアプローチを得ることが出来ようがそれによって後者を直ちに前者の実現形態に還元することは出来ない。

すなわち grades が作用する現実のシステムでは前者の超驗的な演进的展開は尚多くの問題を残している。従ってこゝでは寧ろ両者を明確に切離し、前者を一時的過程的に想定はするが、本来的には後者に忠実に沿った帰納的展開が如何に可能であるか、分配法則の主要性質を如何なる形で確認し得るか、又価格産出量等に認められる不安定性を如何に理解するかが本文の主題となる。問題とするシステムにおける或種の分布法則の成立を価値法則の社会機構を通じての一実現形態であるとするのは当然であるが、その定式化と具体的展開は猶今後の課題である。

### 2. 集中曲線 (concentration curve ; Lorenz diagram) とは competitive system における

普遍的現象の一つである所得の各階層への分配の均等、不均等性を端的にグラフ的に表現したものである。具体的に例えば上位又は下位層からの人質累計百分率を X 軸にその所得又は資産額累計百分率を Y 軸にて表現した曲線である。従ってこの導関数は各 Grade の share を表わしている。

### 3. 集中係数 (concentration coefficient ; variabilita e mutabilita) と平均差 (mean difference)

P. E. Hart によれば上記の system における例えばある二企業の規模の差は他の要因を捨象すれば、price 又は 'output polycy' に関する一方の他方に対する 'the degree of dominance' を与える。

この時この dominance の測度は或産業全体としては、凡ゆる可能な二企業の組合せを考慮して

$$A = \frac{1}{N(N-1)} \sum_{i,j=1,2,\dots,N} |x_i - x_j|$$

即平均差で表すことが出来る。

又、この測度の相対的性格を考慮して

$$G = \frac{A}{2\mu}$$

即 Gimi 係数で表現することが出来る。

### 4. 歪み型分布 [10]

歪み型分布は以上の場合に於て結済量の size distribution に具現されるものであるが、より一般的に云ってピアソン型分布関数論を中心とする統計理論の或種の限界を支える分布と規定出来る。即この分布の代表と云えるパレート型分布ではモーメントの存生は普遍性を持ち得ないし、或場合には中心極限定理の成立すら認められないからである。特にパレート係数  $\alpha$  が  $0 < \alpha < 2$  の範囲に於ては  $\alpha$  は characteristic exponent [10] に一致し、この領域は数理統計学的には殆ど未知の領域といえるからである。

5. 従って 1, 2 の極めて現実的な概念を基礎として 3 に対する解明を試みつゝあるイタリヤ学派を qualitative characteristic [3] の研究組織ということは或意味で妥当性を見出し得る。

## II. 諸方式

## § 1 集中曲線と分布関数

集中曲線は往々分布関数（正確には size distribution）の一属性、一変換様式と考えられているが、実はより包括的現実的概念に立脚している。

然し当初は前者の考えに則してこれを経験的に定義し後により包括的な基礎での定義を試みることにする。

## [I] 集中曲線の定義

## (i) 経験的定義

まず  $N$  箇の互に相なる正值観測値の集団又は系列\*

$$x_1, x_2, \dots, x_N \quad (1)$$

についてこれを定義しよう。

我々は(1)の系列を更にその大きさに従って、小なるものから大なるものに向って並べかえその結果を

$$x_{(1)}, x_{(2)}, \dots, x_{(N)}^{**} \quad (2)$$

で表すことにする。

又此の結果の逐次の累積を

$$S_{(n)} = \sum_{i=1}^n x_{(i)} \quad (3)$$

で表すならば、

**定義 1** 有限系列(1)の集中曲線  $A_n(x)$  とは

$$A_n(x) = \begin{cases} 0 & \text{但し } 0 \leq X < \frac{1}{N} \\ \frac{S_{(n)}}{S_{(N)}} & \text{但し } \frac{n}{N} \leq X < \frac{n+1}{N}, n = 1, 2, \dots, N-1, \\ 1 & \text{但し } X = 1 \end{cases} \quad (4)$$

と  $[0, 1]$  区間で定義される関数である。

もし(1)が  $k (< N)$  個の相異なる観測値

$$x_1', x_2', \dots, x_k' \quad (5)$$

からなり各  $x_i'$  が  $\nu_i$  なる度数をもつ場合には、系列(1)の平均  $\mu_N$  及び  $x_i'$  の相対度数

$$f_i = \frac{\nu_i}{N} \text{ 及び 逐次累積 } S'_{(n)} = \sum_{i=1}^n \nu_i x'_i \quad **$$

を用いて

**定義 2** 有限系列(1)に対する定義1の拡張された集中曲線は

$$A_N(x) = \begin{cases} 0 & \text{但し } 0 \leq X < f_1 \\ \frac{S'_{(n)}}{S'_{(k)}} = \frac{\sum_{i=1}^n \nu_i x'_i}{\sum_{i=1}^k \nu_i x'_i} = \frac{\sum_{i=1}^j f_i x'_i}{\mu_N} & \text{但し } \sum_{i=1}^n f_i \leq X < \sum_{i=1}^{j+1} f_i; j = 1, 2, \dots, k-1 \\ 1 & \text{但し } X = 1 \end{cases} \quad (6)$$

\* 観測値と云うがここでは観測行為とか、それに伴う変動を考慮せず単に確定した数量の集りを指している。単一の客体の得返し観測によって得られたものであるが又集団の客体の同時観測によるか等も問題にしない。

\*\* ( ) は一般に順序づけられた系列に対する場合を示すことにする。

(4) 及び (6) は階段型関数であるが、場合によってはこの各不連続点を折線で結び連続化した集中折線を用いるのが都合がよい\*。

扱我々は定義(2)を更に押し進めて、一般に正值観測値の無限系列

$$x_1, x_2, \dots, x_N, \dots \quad (7)$$

についての集中曲線を定義する。

**定義 3** 無限正值観測値系列 (7) の集中曲線とは

この系列の第 1 項から第  $N$  項までの部分系列に関する集中曲線の極限関数が存在するときこの関数で定義する。

$$\text{即 } A(X) = \lim_{N \rightarrow \infty} A_N(X), \quad 0 \leq X \leq 1 \quad (8)$$

である\*\*。

例えば (7) の系列が

$$1, 2, \dots, N, \dots$$

で与えられている時

$$A_N(x) = \begin{cases} 0 & 0 \leq X < \frac{1}{N} \\ \frac{n(n+1)}{N(N+1)} & \frac{n}{N} \leq X < \frac{n+1}{N}, \quad n = 1, 2, \dots, N-1 \\ 1 & X = 1 \end{cases}$$

従って定義 3 により

$$A(X) = X^2 \quad \text{又} \quad \mu = \lim_{N \rightarrow \infty} \frac{N+1}{N} = \infty$$

ここで系列の考えを離れ、観測の基礎に連続な変量が考えられ、それに対応する密度関数  $f(x)$  が又連続である時 (6) 及び (8) の形式に対応して次の定義が可能である。

**定義 4** 区間  $0 \leq x \leq \infty$  で定義される連続な密度関数に対し、その集中曲線は

$$X = \int_0^x f(\xi) d\xi$$

$$A(X) = \frac{1}{\mu} \int_0^x \xi f(\xi) d\xi \quad (9)$$

で定義される。但し  $\mu = \int_0^x \xi f(\xi) d\xi < \infty$  とする。

この場合  $A(X)$  は微分可能であり、容易に次式を得ることが出来る\*\*\*。

$$\begin{cases} A'(X) = \frac{x}{\mu} \\ A''(X) = \frac{1}{\mu f(x)} \end{cases} \quad (10)$$

(6) 及び (8) で定義されるような discrete な変量に対して集中折線は右方微係数の階差を  $\Delta^2 A(X)$  で表せば

$$\begin{cases} D^+ A(X) > 0 \\ \Delta^2 A(X) \equiv D^+ A(X) - D^+ A(X') \geq 0 \quad X > X' \end{cases} \quad (11)$$

この (10) 式の第一式から分るように、 $A'(X)$  は、観測個体の順位づけられた share distribution を示すものであり、又 (10) 式の第二式で明瞭なように、 $A''(X)$  は size distribution

\* 猶大ききの大なる順に累積した集中折線は  $(\frac{1}{2}, \frac{1}{2})$  を中心にして  $180^\circ$  回転した形になる。

\*\* この中合  $\mu_N$  の極限  $\mu$  の存在又は  $\mu$  が 0 でない等は限らずしも必要でない。[11]

\*\*\* 逆は必ずしも云ない。[11]

そのものの特質を示している。

従って集中曲線は、この二つの distribution の統一形式といえるわけである。

以上で集中曲線は凡て正值変量について考えられたが此の拡張は極めて容易であることを附記しておく。

(ii) 集中曲線の独自の基礎づけ

密度関数を用いた場合集中曲線は以上のように定義され得るが、序文、脚注\* 及び本文中における各種の集中曲線の特性に照してみるならば、此の曲線は独自の性格を持ち、密度関数の存在を必ずしも前提する必要がないことが分る。

又集中曲線の実際上の適用領用を考えるならば、単に確率論的に規定された密度関数のみに立脚すきものではなく、直接現実的諸概念とその関係の定式化、公理化の上に定義づけらるべきことは略明かとなる。

即 (i) 母胎となる価値所有者層を数学的抽象空間に於て表現し

(ii) 逆に価値所有者層をこの空間に含まれる集合として、空間、集合の数学的規定を与え

(iii) 此等所有者層の所有する価値量を集合の上に定義される集合関数として捉え

(その際全空間に対しての相対的価値量 (share) に対応してこの集合関数は、価値の分配 ( $D$ ) の原則に則応して公理化されることになる

(iv) 更に所有者の grading を上記の空間上の点の  $R_1$  空間への写像 (正確には半空間  $(0 \leq \xi \leq \infty)$  への写像を ranking,  $0 \leq \xi \leq 1$  への写像を grading とすることができ) とするならば、

系列曲線は

$$M(X) = D(S_x)$$

であり、こゝで  $S_x$  は grade  $\xi \leq X$  に対応する集合、 $D$  はこの層への分配量を示す。

特定な grading の下では  $M(X)$  は  $X$  軸に対し凸となるが、此の種の grading が存在するときそれに対する  $M(X)$  を特に集中曲線  $A(X)$  とする。又かゝる grading の存在について、確率論が有効な条件を与えている。

然し以上は猶不充分であり、此の詳細な記述、及び論理的数学的展開は次の段階に於ける課題である。

つまり本稿の目的は (i) の諸経験的定義に出発して、(ii) の本質的規定による帰結の一端を予想することにあるからである。

集中曲線の実例

1. Pareto 型分布

第1表参照

2. truncated Cauchy 分布

$$f(x) = \frac{2}{\pi} \frac{a}{1+x^2} ; k > x > 0, k = \tan \frac{\pi}{2a}$$

とすると

$$X = F(x) = \frac{2}{\pi} a \tan^{-1} x$$

$$Y = \frac{1}{\mu} \int_0^x \xi dF(\xi) = \frac{a}{\pi \mu} \log(1+x^2)$$

$$\text{こゝに } \mu = \int_0^k \xi dF(\xi) = \frac{2a}{\pi} \log \sec \frac{\pi}{2a}$$

$$\text{従って } Y = \frac{a}{\pi \mu} \log \sec^2 \frac{\pi}{2} X = \frac{\log \sec \frac{\pi}{2a} X}{\log \sec \frac{\pi}{2a}}$$

パレート型分布族  
第1表 集中曲線の分類

Types	$k$	集	中	Limit curve
(A)	$k < 0$	$Y = \frac{(1+\gamma X)^\beta - 1}{(1+\gamma)^\beta - 1}$ , ここで $1 < \beta < 2$ 且 $\gamma > 0^*$		$Y = X^\beta$ ( $a \rightarrow 0^*$ )
(B)	$k = 0$	(i) $Y = X^2$ ( $A \leq 1$ ) (ii) $Y = X$ ( $A = 1$ )		
(C)	$0 < k < 1$	$Y = \frac{(1+\gamma X)^\beta - 1}{(1+\gamma)^\beta - 1}$ , ここで $2 < \beta \leq \infty$ 且 $\gamma > 0^*$		$Y = X^\beta$ ( $a \rightarrow 0^*$ )
(D)	$k = 1$	$Y = \frac{e^{\beta X} - 1}{e^\beta - 1}$ , $\beta > 0$	(Zipf)	
(E)	$1 < k < 2$	$Y = \frac{(1+\gamma X)^\beta - 1}{(1+\gamma)^\beta - 1}$ , ここで $-\infty \leq \beta < 0$ 且 * $-1 < \gamma < 0$ (Mandelbrot)		
(F)	$k = 2$	$Y = \frac{\log(1+\gamma X)}{\log(1+\gamma)}$ , ここで $-1 < \gamma < 0$		
(G)	$k > 2$	$Y = \frac{(1+\gamma X)^\beta - 1}{(1+\gamma)^\beta - 1}$ , ここで $0 < \beta < 1$ 且 $-1 < \gamma < 0^*$		$1 - Y = (1 - X)^\beta$ ( $b \rightarrow \infty$ ) * (Gini's $\delta$ -curve)

\*  $\beta$  は一般に (A), (C), (E) (G) に於ては  $(k-2)/(k-1)$  に等しい。

### 3. 指数分布

$$f(x) = \lambda e^{-\lambda x} ; \quad x > 0 \text{ とすると}$$

$$X = F(x) = 1 - e^{-\lambda x}$$

$$Y = \frac{1}{\mu} \int_0^x \xi f(\xi) d\xi$$

$$= \lambda \left( -x e^{-\lambda x} + \frac{1 - e^{-\lambda x}}{\lambda} \right) = (1 - X) \log(1 - X) + X$$

#### 集中曲線の主要性質

[I] で定義された集中曲線は当面次のように分けて考えられる。

- (1) 微分可能なもの (連続な密度関数に対応する場合等)
- (2) 微分可能ではないが連続なもの (集中折線に対応するもの等)
- (3) 非連続なもの (有限観測値に関する段階型のもの等)

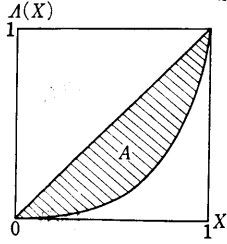
又共通した性質として次のものが挙げられる。

- (1)  $A(0) = 0$ ,  $A(1) = 1$  且  $0 < X < 1$  に対し  $0 \leq A(X) \leq 1$
- (2)  $A(X)$  は非減少関数である
- (3)  $A(X)$  は至る処右連続である
- (4)  $A(X)$  は可附番個の点を除き微分可能である
- (5) もし  $A(X)$  が連続ならば上に凹である

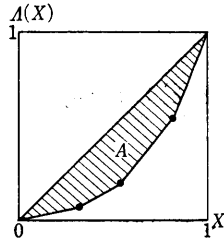
従って第 1, 2, 3 図の如く表現さる。

扱、此の曲線について、我々が分布関数の考察の時考えたように、記述的尺度を入れて更に具体的に考察することにする。

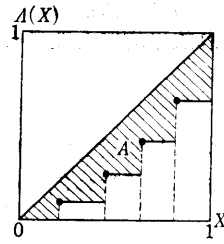
以不に於て我々は屢々  $X$  を grade 又は fraction,  $\mu A'(X)$  を share 又は fractile と云う場合がある。又  $A(X) = X$  で示される対角線を均等線という。



第1図 微分可能な集中曲線



第2図 連続型の集中折線



第3図 階段型の集中曲線

(i) 代表値

定義5 平均値に当る grade  $X_\mu$  を集中曲線の代表値とする。

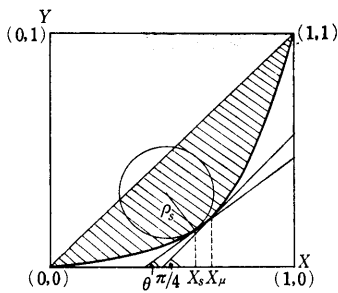
微分可能な集中曲線については

$$A'(X_\mu) = 1 \tag{12}$$

が成立する。

(ii) 集中度

定義5 集中曲線の集中度比率 (ジニ係数)  $G$  によって与えることができる。(1. 基礎概念及び第4図集中曲線の記述的測度 (巻末) 参照)



第4図 集中曲線の記述的測度

即, 例えば系列 (1) に於て平均差

$$A = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=2, \dots, N} |x_i - x_j| \tag{13}$$

を用いて

$$G = \frac{A}{2\mu} \tag{14}$$

によって定義可能である。連続な密度関数をもつ集中曲線については

$$A = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - y| f(x) f(y) dx dy \tag{15}$$

で与えられる。

微分可能型の集中曲線 (第1図) に対し,  $f(x)$  が連続の場合には (15) 式を用いて

$$G = 2A \tag{16}$$

が容易に得られる。

又第2図の如き折線に於ては (13) 式を用いて

$$\frac{N-1}{N} G = 2A \tag{17}$$

が成立する。

従ってそれぞれの場合について

$$G = 1 - 2 \int_0^1 A(X) dx \tag{18}$$

$$G = 1 - 2 \frac{1}{N-1} \sum_{n=1}^{N-1} \frac{S_{(n)}}{S_{(N)}} \tag{19}$$

等が成立する。

故に

$$0 \leq G \leq 1 \tag{20}$$

又後の場合

$$0 \leq G \leq \frac{N}{N-1} \tag{21}$$

が云えることになる。

それ故平均差を改良して

$$d' = \frac{1}{N^2} \sum_{i,j=1,2,\dots,N} |x_i - x_j| \quad (22)$$

を用いるならば常に (20) が保たれる。

この集中係数の代りに集中度として

$$D = 2 \{A(x_\mu) - X_\mu\} \quad (23)$$

を用いることが出来るが、これは平均偏差

$$\delta = \int_{-\infty}^{\infty} |x - \mu| f(x) dx \quad \text{又は} \quad \delta = \frac{1}{N} \sum |x_i - \mu_N|$$

を用いると

$$D = \frac{\delta}{\mu} \quad (24)$$

であることが分る。

平均差、平均偏差及び標準偏差については次のような関係等が知られている。

$$\frac{N}{N-1} \leq \frac{d}{\delta} \leq 2 \quad (\text{Glasser の不等式})$$

$$\begin{aligned} (2/N)^{1/2} = \frac{\delta}{\sigma} &\leq (1 - N^{-2})^{1/2} && N \text{ 奇数 (Chakravarty の不等式)} \\ &\leq 1 && N \text{ 偶数} \end{aligned}$$

集中係数の実例

### 1. Pareto 型分布

$$f(x; k, A) = \frac{A}{x^k} \quad (x > 0)$$

第 3 表参照

### 2. 指数型分布

$$f(x; \lambda) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \quad (x > 0)$$

$$G = 1 - 2 \int_0^1 A(x) dx = \frac{1}{2}$$

常に一定である。

### 3. Gibrat 分布

$$f(x; \theta, \lambda) = \frac{1}{x \lambda \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log x - \theta}{\lambda} \right)^2} dx$$

$\varphi(x)$  を規準正規分布とすれば

$$G = 1 - 2 \varphi \left( -\frac{\lambda}{\sqrt{2}} \right) = 2 \varphi \left( \frac{\lambda}{\sqrt{2}} \right) - 1$$

等

### (iii) 歪 度 [10]

集中曲線に於ては (1, 0) (0, 1) を結ぶ対角線について対称な形が基準と考えられる。この性質は例えば Gibrat 分布に於て見出すことが出来る。

従って歪度として次の定義を与えることが出来る。

**定義 6** 微分可能な集中曲線に於て曲率最大を与える点の grade を  $X_s$  とするとき、又この点に於ける接線が  $X$  軸と交わる角度を  $\theta_s$  とするとき、 $S = \cos 2\theta$  をこの曲線の歪度とする。

$X_s$  に対応する変量を  $x_s$  (これを飽和値とする) とすれば  $S$  は又



$$S = \frac{\mu^2 - x_S^2}{\mu^2 + x_S^2} \quad (25)$$

で与えられ

$$-1 \leq S \leq 1 \quad (26)$$

が明かである.

猶連続な密度関数が自己対称である必要充分条件は  $\mu$  を平均値とした場合

$$\frac{f\left(\frac{\mu^2}{x}\right)}{f(x)} = \left(\frac{x}{\mu}\right)^3 \quad (27)$$

が成立することであり, 連続な二つの密度関数  $f_1(x)$ ,  $f_2(x)$  が相互に対称である条件は

$$\frac{\mu_2 f_2(x')}{\mu_1 f_1(x)} = \left(\frac{x}{\mu_1}\right)^3 \quad (28)$$

こゝに  $\mu_2 x' = \mu_1 x$ ,  
である.

称対集中曲線の実例

### 1. Pereto 型分布

後述 (IV) の如くパレート分布ではパレート恒数  $\alpha$  につき  $\alpha = 1.5$  のとき  $S = 0$  自即自己対称であり  $\alpha = 1.5 + p$  と  $\alpha' = 1.5 - p$  とでは  $S_\alpha = -S_{\alpha'}$  であり相互対称である.

### 2. 自由度2の $t$ 分布

集中曲線は  $(0, 0)$   $(1, 1)$  を通る円弧である.

### 3. Gibrat 分布は帰述のように常に自己対称である.

(v) 尖 度 [10]

**定義 7** 集中曲線の尖度  $K$  は  $X_S$  に於ける曲率半径  $\rho_S$  で与えることが出来る.

$$\rho_S \text{ はまた } \rho_S = \min \frac{\left(1 + \frac{x^2}{\mu^2}\right)^{3/2}}{\mu f(x)} \quad (29)$$

で与えられる.

もし,  $R$  をこの集中曲線と同じ弧長をもち, 点  $(0, 0)$  及び  $(1, 1)$  を通る円の半径とすれば

$$0 \leq 1 - \frac{\rho_S}{R} \leq 1$$

が明であり, 従ってこれを標準化された測度にすることが出来る.

今集中曲線の最大曲率を尖度  $K = \frac{1}{\rho_S}$  とした時 non-skew 分布は  $K = \infty$  であり, それは  $X_S = 1$  に於てである.

逆に skew 分布は  $K < \infty$  であるか, 又は  $X_S = 0$  である.

従ってこの測度は分布が歪み型であるかを判別する一方, 歪み型分布に対する集中曲線の有効性を与えている.

## III. 集中曲線の間関係

今  $x$  及び  $y$  に関する集中曲線をそれぞれ  $A(X; x)$ ,  $A(X; y)$  とすれば一般に次の関係が認められる.

(i)  $y = x + a$

の場合, 任意の  $X$  に対して

$$\frac{X - A(X; y)}{X - A(X; x)} = \frac{\mu_y}{\mu_x} = \frac{G_y}{G_x} \quad (30)$$

が成立する.

(ii)  $y = \lambda x$  の場合

$$A(X; y) = A(X; x) \quad (31)$$

(iii)  $y = x^p$  の場合両者の密度を  $f(x)f(y)$  とすれば

$$f(y) = \frac{f(x)}{p x^{p-1}}$$

従って  $f(x)$  がパレート型分布 (後述 [IV])

$$f(x) = \frac{A}{x^k}$$

ならば

$$f(y) = \frac{B}{x^{k+p-1}}$$

即歪度  $K_x, K_y$  につき

$$K_y - K_x = \frac{2}{3} (1 - p) \quad (32)$$

だけ歪みの差をもつ。

又  $f(x)$  が Gibrat 仮説

$$\Phi\left(\frac{\log x - \theta}{\lambda}\right)$$

に従うならば前述のように Gimi 係数は

$$G_x = 2\Phi\left(\frac{\lambda}{\sqrt{2}}\right) - 1$$

$$G_y = 2\Phi\left(\frac{\lambda}{\sqrt{2}p}\right) - 1$$

と変化する。

集中曲線の変化の実例

1. 基準時点における  $x$  の集中曲線  $A_0(x)$  と  $t$  時刻の集中曲線  $A_t(X)$  の間には上の (i) (ii) (iii) に対応して、均等成長、比例成長、弾力的成長が考えられ、それぞれの場合につき

$$(i) \quad \frac{X - A_t(X)}{X - A_0(X)} = \frac{\mu_t}{\mu_0} = \frac{G_t}{G_0}$$

$$(ii) \quad A_t(X) = A_0(X)$$

$$(iii) \quad K_t - K_0 = \frac{2}{3}(1 - \beta) \quad \beta \text{ は弾力性係数}$$

が成立する。

2. 1 の具体例は亦、大衆課税、所得に比例した課税、累進課税の所得分配の直接的効果のモデルを与えるであろう。
3. 又一方で 1 の例は例えば Cobb-Douglas 型の生産函数を仮定した場合資本係数、資本集約度等が経営規模にどのような関係をもつかによって資産分布のパレート係数、ジブラ係数等の変化を予測する根拠となる。

#### IV. 歪み型分布関数族の集中曲線群

こゝでは集中曲線の展型的な例題を与えることにする。

序文にも触れた歪み型分布族は competitive system における経済量に屢々認められているが、その分布関数としての特徴付けは必ずしも一致していない。そのうちでもごく代表的と思われる定義としては

$$f(x) \sim x^{-k} \quad (x \rightarrow \infty); \quad 0 \leq k \leq 3 \quad (33)$$

を挙げることが出来る。但し、例えば所得分布の一形式とされている Gibrat 分布はこの定義に従わないが、これを歪み形分布として扱うものも多い。又場合により  $-\infty < k < \infty$  も含めると扱い易い。

此の定義の範囲内にあるものは Pareto, Zipf, Yule, Levy-Pareto 等が知られているが、こゝでは定義の形からしても最も代表的と思われる Pareto 分布

$$F(x) = 1 - \left(\frac{\theta}{x}\right)^\alpha, \quad 0 < \theta \leq x, \alpha > 0 \quad (34)$$

或は密度関数の形で

$$f(x) = \frac{\alpha \theta^\alpha}{x^{1+\alpha}}$$

を取挙げてみると次のように著しい特徴をもつことが容易に分る。

- (i)  $\alpha \leq 1$  の場合に平均値は存在しない。  
 $\alpha \leq 2$  に対しては分散が存在しない。  
 一般に  $\alpha \leq n$  に対し  $n$  次のモメントは存在しない。
- (ii)  $0 < \alpha < 2$  に対し中心極限定理は成立しない。この  $\alpha$  は characteristic exponent と呼ばれる値に一致するものである。
- (iii) 此の分布の幾何平均は  $\theta e^{1/\alpha}$  で与えられる。

以上で平均、分散或は一般に  $n$  次モメントを用いて分布関数の特徴づけ或は統計的推論に於て数理統計学的方法がそのまづでは適用困難であることを知る事が出来るであろう。

翻って以上の事柄を [I] [II] で定義された集中曲線によって特徴づけてみよう。

先ず (33) に与えた歪み分布族とは  $X = 1$  で飽和値を持たぬ場合であり、もし持ったとしても曲率で定義された集中曲線の尖度  $K$  が有限である場合である。逆に歪み分布族以外では  $X = 1$  で飽和値をもち尖度  $K$  はこの点の曲率で与えられ且無限大となる。

次に (34) の Pareto 分布については

- (i)  $\alpha = 0.5$  に於て集中曲線は自己対象曲線である。  $\alpha = 0.5 + p$  と  $\alpha = 0.5 - p$  とをもつ二つの集中曲線は相互対称である。この関係は有限区間  $[a, b]$  , 且  $k \geq 0$  の分布

$$f(x) = \frac{A}{x^k}; \quad a \leq x \leq \quad (35)$$

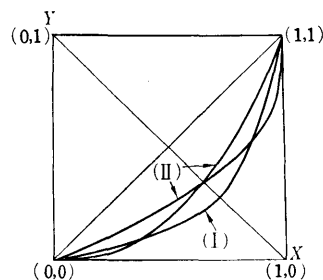
についても成立するから、集中曲線はあながちに歪み型分布族についてのみ有効なわけではない。

- (ii) (35) は  $k > 1$  の場合 (34) の切断された分布を示しているが、一般に (34) の曲線は  $\theta$  の値に対し不変であり (35) の曲線は  $a$  の値に対し不変である。

次に (34) (35) の集中曲線を [I] の定義に従って具体的に表現することにする。第 1 表及び第 2 表及び第 6 図) はその結果であるが、その結果 (34) (35) の分布は集中曲線の特徴から細分類される事が分る。

## V. データ解析の実際例

では実際に日本の経済統計資料から得られた結果はどうなるであろうか。資料は聯か遡るが全数に近い標本で且色々な意味で基準年次とされる時点に行われたという点で昭和 30 年国富調査を例にとってみよう。第 8, 9, 10, 11 図

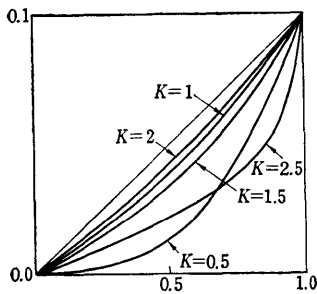


第 5 図 自己対称型 (I) 及び相互対称型 (II) 集中曲線

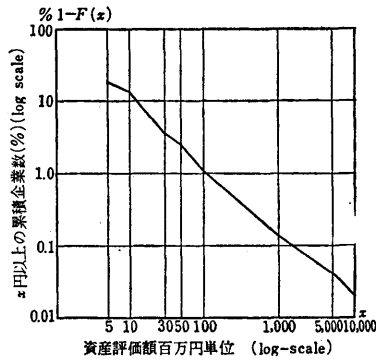
第 2 表 第一表におけるパラメータの性質

Types	Parameters		Limit curve
	$\beta$	$\gamma$	
(A)	$\beta > \frac{b}{\mu} *$	$\frac{F(a)}{1-F(a)}$	$\beta = \frac{b}{\mu} \quad (a \rightarrow 0)$
(B)	$\beta = \frac{r}{\mu}$ for $A \neq 1$		
(C)	$\frac{b}{\mu F(a)} > \beta > \frac{b}{\mu} *$	$\frac{F(a)}{1-F(a)}$	$\beta = \frac{b}{\mu} \quad (a \rightarrow 0)$
(D)	$\beta = \frac{r}{\mu}$		
(E)	$-\frac{r}{\mu F(b)} < \beta *$		
(F)			
(G)	$\beta < \frac{a}{\mu F(b)} *$	$-F(b)$	$= \frac{a}{\mu} \quad (b \rightarrow \infty)$

\*  $\beta$  は一般に (A), (C), (E), (G) に於ては  $r+b\gamma/\mu\gamma$  に等しい。



第 6 図 パレートの型分布  
( $-\infty < K < +\infty$ )

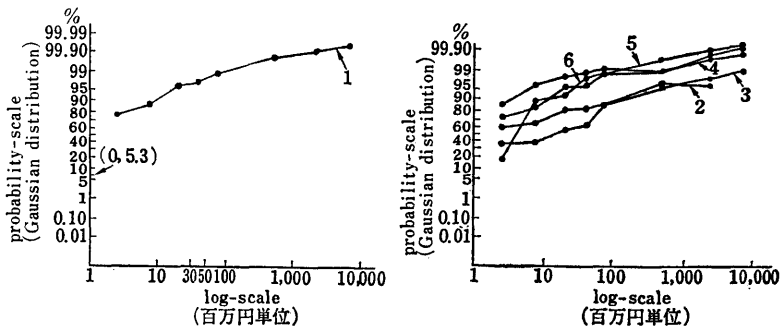


第 7 図 全産業に於ける資産の  
sizedistribution

$$-F(x) = \left(\frac{\theta}{x}\right)^\alpha \text{ 且}$$

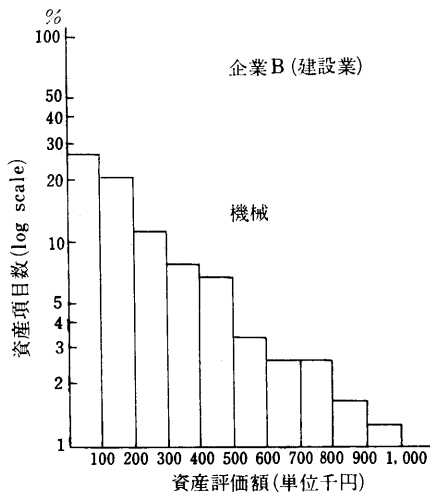
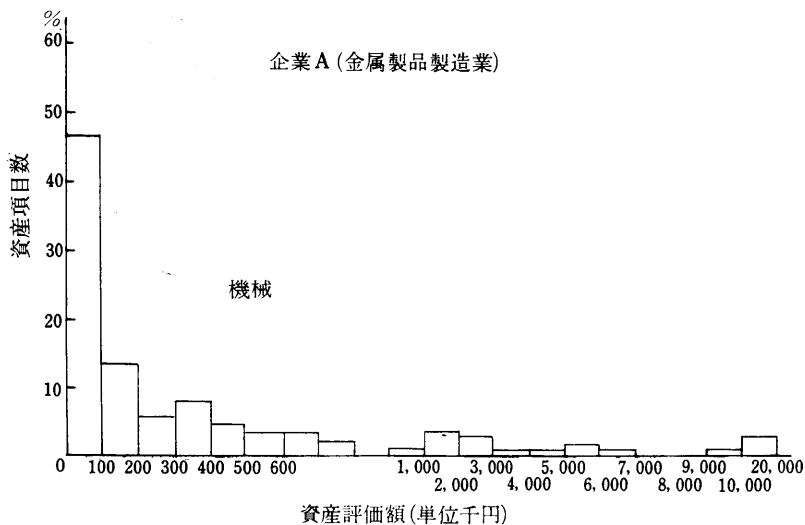
$$\alpha = \frac{1}{2G} + \frac{1}{2} = 1.06517$$

(Table 8 参照)



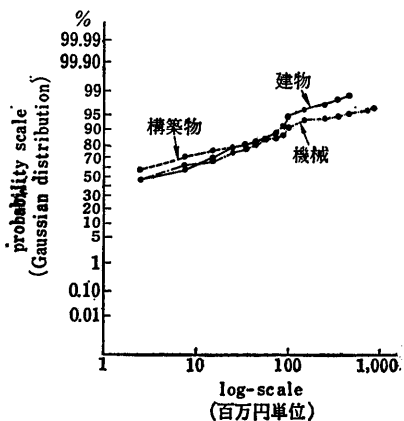
- 1. 全産業 2. 金融業 3. 製造業 4. 水産業
- 5. 建設業 6. 金融、保険業

第 8 図 特定産業における資産の Size distribution

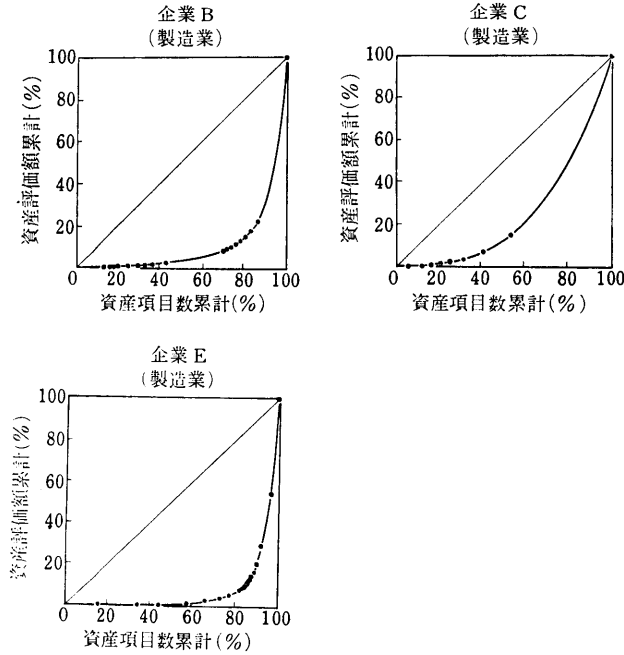


第9図 特定企業における資産分布\*

\* データ源；1955年（昭和30年）度日本における国富調査に対する試験調査結果



第10図 某企業における資産の size distribution



第 11 図 特定企業における資産の集中曲線

及び第 5 表は実際に以上に基づいて算出した結果であるが、此等に共通して云える事は [II] の諸測度について

$$\frac{3}{4} < X_\mu < 1, \quad \frac{1}{3} < G < 1, \quad -1 < S < -\frac{1}{3} \tag{36}$$

を示すということである。

これ等の結果から (34) 式の  $\alpha$  を推定してみると第 1 表の (G) タイプにもれなく該当するのであり、此のことは国際的にみても例外をなすものではない。第 12 図は推定値  $\alpha$  に基づく理論曲線との比較を示している。

この差異は  $\alpha$  の算出が第 1 表 (G) の極限形式によったからであり、最大規模によって決定される  $\gamma$  を入れれば可成り適合度は高まるであろう。

### § 2 系列曲線と系列の性質

§ 1 の集中曲線の概念を更に一般化することを考えてみよう。

§ 1 では (1) の観測値系列から (2) を導いたが

$$S_n = \sum_{i=1}^n x_i \tag{37}$$

を用いて直接 (1) について一つの関数を定義する。

**定義 8** 有限系列の系列曲線  $M_N(X)$  とは

$$M_N(X) = \begin{cases} 0 & \text{ここで } 0 \leq X < \frac{1}{N} \\ \frac{S_n}{S_N} & \text{ここで } \frac{n}{N} \leq X < \frac{n+1}{N}; n = 1, 2, \dots, N-1 \\ 1 & \text{ここで } X = 1 \end{cases} \tag{38}$$

で与えられる rath である。

\* R. C. Bliz 及び J. A. Brittain (1964) [2] は此の曲線を相関曲線 correlation curve と名づけた我々は S3 との関係で名称を変更した。

第3表 Gini 係数の  $k$  に関する変動

Types	$k$	$G$	Remarks	Limit curve
(A)	$k < 0$	$< \frac{\beta-1}{\beta+1} \left( = \frac{1}{3-2k} \right)^*$	$G$ は $k$ と共に増大する	$G = \frac{\beta-1}{\beta+1} (a \rightarrow 0)$
(B)	$k=0$	(i) $\frac{1}{3}$ for $A \geq 1$ (ii) $0$ for $A=1$		
(C)	$0 < k < 1$	$< \frac{\beta-1}{\beta+1}$	$G$ は $k$ と共に増大する	$G = \frac{\beta-1}{\beta+1} (a \rightarrow 0)$
(D)	$k=1$	$1 - \frac{2}{\beta} + \frac{2}{e^3-1}$ $1 - \frac{2}{\beta} < G < -1 - \frac{2}{\gamma}$	$G$ は $\beta$ と共に増大する	
(E)	$1 < k < 2$	(ii) $1 - 2 \frac{\frac{\log(1+\gamma)-1}{\gamma}}{(1+\gamma)^{-1}-1}$ , for $k=1.5^*$	$G$ は $\gamma$ が減力するとき増大する	
(F)	$k=2$	$-1 - \frac{2}{\gamma} + \frac{2}{\log(1+\gamma)}$	$G$ は $\gamma$ が減力するとき増大する $G=0 (\gamma \rightarrow 0)$ 又 $G=1 (\gamma \rightarrow -1)$	
(G)	$k > 2$	$< \frac{-\beta+1}{\beta+1}$	$G$ は $k$ が増大するとき減少する	$G = \frac{-\beta+1}{\beta+1} (b \rightarrow \infty)$

\* 一般に  $G = 1 - 2 \frac{(1+\gamma)^{\beta+1} - 1}{\gamma(\beta+1)(1+\gamma)^{\beta-1}}$  が (A), (C) (E) 及び (但し  $k=1.5$  を除く) (G) に於て成立する

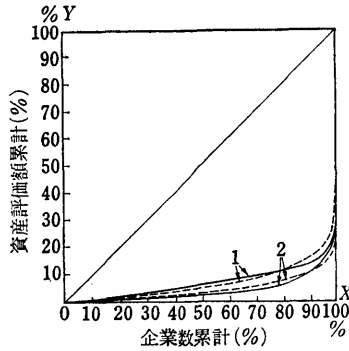
第4表  $k$  に関する曲線の歪みの変動

Types	$k$	$S$	Saturation value (飽和値)	Remarks
(A)	$k < 0$	$\frac{1-\beta^2}{1+\beta^2}$	$x_S = \mu \beta = b (a \rightarrow 0)$	
(B)	$k=0$	1	存在せず	
(C)~(F)	$0 < k < 2$	$1 - \frac{2}{3} k$	$\frac{\mu}{\sqrt{\frac{3}{k}-1}}$	$-1 < S < 0$ ( $0 < k < 1.5$ ) $S=0$ ( $k=1.5$ ) $0 < S < 1$ ( $k > 1.5$ )
(G)	(i) $2 \leq k < 3$ (ii) $k \geq 3$	$1 - \frac{2}{3} k$ $\frac{1-\beta^2}{1+\beta^2}$	$\frac{\mu}{\sqrt{\frac{3}{k}-1}}$ $x_S = \mu \beta = a (b \rightarrow \infty)$	$0 < S \leq 1$

\* (C)~(F) 及び (G) (i) における  $S$  の式は  $[-\infty, \infty]$  にまで拡張され、この時極限集中曲線に於ては  $S=3/G$  に等しい。

集中曲線の場合と同様にこゝでも系列折線をも考えることが出来る。

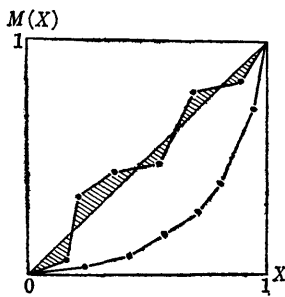
この曲線も正值系列に対して単調増大であるが上に凹であるとは云えない。集中曲線は第13図に示すようにこの曲線の限界を与える。



1. 建設業  
 2. 製造業  
 — 経験的曲線  
 .....  $\beta = \frac{1-G}{1+G}$  を用いた Gini's  $\delta$  曲線  
 即 Table 6(G) (limit curve) 型の理論曲線

第 12 図 理論的曲線と経験的曲線

又この場合も均等線との間の面積  $A$  (但し均等線の下部は正, 上部は負の値をとる) は



第 13 図 系列折線

$$2 A_N = \frac{\frac{1}{N^2} \sum_{i < j} \sum_{i, j=1, 2, \dots, N} (x_i - x_j)}{\mu_N}$$

$$= \frac{\frac{1}{N^2} \sum_{i > j} \sum_{i, j=1, 2, \dots, N} \text{sgn}(x_i, x_j) |x_i - x_j|}{\mu_N}$$

$$= \frac{\frac{1}{N^2} \sum_{i, j=1, 2, \dots, N} \sum_{i, j=1, 2, \dots, N} \text{sgn}(i, j) (x_i - x_j)}{2 \mu_N} \quad (39)$$

ここで

$$\text{Sgn}(x, y) = \begin{cases} -1 & x < y \\ 0 & x = y \\ +1 & x > y \end{cases} \quad (40)$$

今  $2 A_N = C_N$  (41)

又  $\left. \begin{aligned} \frac{1}{N(N-1)} \sum_{i, j=1, 2, \dots, N} \sum_{i, j=1, 2, \dots, N} \text{sgn}(i, j) (x_i - x_j) &= \delta_N = \delta_{x, N} \\ \frac{1}{N^2} \sum_{i, j=1, 2, \dots, N} \sum_{i, j=1, 2, \dots, N} \text{sgn}(i, j) (x_i - x_j) &= \delta'_N = \delta'_{xN} \end{aligned} \right\} \quad (42)$

としておく。

(i) この時  $\max C_N = G_N$  及び  $\max \delta_N = \Delta_N$  且  $-1 \leq C_N \leq 1$  である。  
 又 Kendall の順位相関係数  $\tau$  は

$$\tau = \frac{2 \sum_{i > j} \sum_{i, j=1, 2, \dots, N} \text{sgn}(x_i, x_j)}{N(N-1)} \quad (43)$$

であるから

$$C_N = \frac{C_N}{G_N} = \frac{\delta_N}{\Delta_N} \quad (44)$$

は,  $\tau$  に  $x$  の大きさの weight をつけた形で順位相関係数の機能を果している。

(ii) 今  $x = x_u$  が区間  $[0, 1]$  で  $u$  の連続な正值関数であるとすれば系列曲線は

$$\begin{cases} X = t \\ M(X) = \frac{\int_0^t x_u du}{\int_0^1 x_u du} \quad 0 \leq t \leq 1 \end{cases} \quad (45)$$



と定義することが出来る。

従って

$$M'(X) = \frac{x_t}{\mu_x} \quad . \quad \text{ここで} \quad \mu_x = \int_0^1 x_u du \quad (46)$$

$$\begin{aligned} \text{又} \quad C &= \int_0^1 X dM(X) - \int_0^1 M(X) dX \\ &= \frac{1}{\mu_x} \int_0^1 t x_t dt - \frac{1}{\mu_x} \int_0^1 \int_0^1 x_u du dt = \frac{1}{\mu_x} \int_0^1 \int_0^t (x_t - x_u) du dt \end{aligned}$$

一方

$$\int_0^1 \int_0^1 x_t du dt = \int_0^1 \int_0^1 x_u du dt = \mu_x$$

であるから、結局

$$C = \frac{1}{2\mu_x} \int_0^1 \int_0^1 \text{sgn}(t, u) (x_t - x_u) du dt \quad (47)$$

を得る。ここで  $\text{sgn}(t, u)$  は (40) の関数である。従って  $x$  が  $u$  に関し一定であれば、それは  $(0, 0)$ , (11) を結ぶ対角線に対応する。

他方  $x$  が更に  $u$  の単調増加関数であるとすれば (45) 式は  $x$  に関する集中曲線を与える。

従って集中曲線はもし  $x$  の分布が一定であるとすれば任意の連続な変換  $x_u$  に対する系列曲線  $M(X)$  の下限を示し  $\delta$  係数は (47) 式により  $C$  の上限を与える。

同様に  $x$  が  $t$  の単調減少関数ならば  $M(X)$  の上限及び  $C$  の下限を与える逆順位の集中曲線を得る。

(42) に対応して又

$$\delta_x = \int_0^1 \int_0^1 \text{sgn}(t, u) (x_t - x_u) du dt \quad (48)$$

が定義される。

(iii) 再び (i) の系列及び系列曲線を考えるならば  $\delta_N$  は系列 (1) の変化に従って変化することが分る。即

(A)  $x_k$ ;  $1 \leq k \leq N$  が  $x'_k = x_k + \Delta x$  に変化すると

$$\delta'_{x',N} - \delta_{x,N} = \frac{2(2k - N)}{N^2} \Delta x$$

又 (B)  $h$  番目と  $k$  番目  $1 \leq h, k \leq N$  が入れ換わると

$$\delta'_{x',N} - \delta_{x,N} = -\frac{4}{N^2} (h - k) (x_h - x_k)$$

(C) 系列 (1) に個体  $x_{N+1}$  が追加された場合

$$(N+1)^2 \delta'_{N+1} = N^2 \delta'_N + 2N(x_{N+1} - \mu_N)$$

の如く変化する量である。

以上から  $\delta'_N$  の変化も容易に知ることが出来る。

$\delta'$  の変動の実例

$N$  個の個体からなる集団に於ては上位 1 位及び 2 位の各社が単純に合併したとすれば、まず (C) により、1 位を除いたとした時の変化

$$(N-1)^2 \delta'_{N-1} = N^2 \delta'_{x,N} - 2(N-1)(x_N - \mu_{N-1})$$

を得、更に (A) に於て  $k = N-1$ ,  $\Delta x = x_N$  とすれば

$$\delta'_{x',N-1} = \delta'_{x,N} + \frac{2\{2(N-1) - (N-1)\}}{(N-1)^2} x_N$$

となる。従って合併後の値は両式より

$$\delta'_{x',N-1} = \frac{N^2}{(N-1)^2} \delta'_{x,N} + 2 \frac{x_1 + x_2 + \dots + x_{N-1}}{(N-1)^4}$$

平均差は明に上式の  $x_1 \cdots x_{N-1}$  を順位統計量と考えればよい。従ってこの効果は

$$\delta'_{x',N-1} - \delta'_{x',N} = \frac{2N-1}{(N-1)^2} \delta'_{x',N} + 2 \frac{x_1 + \cdots + x_{N-1}}{(N-1)^4} > 0$$

である。

(iv) 然し  $\delta'_{x',N}$  は次の点に関しては一貫した生質を示している。

(A)  $x_i$  が trend をもつ場合 即

$$x_i = a_0 + a_1 t + \cdots + a_k t^k \quad 1 \leq k \leq \infty \tag{49}$$

の時  $C_N \rightarrow C (N \rightarrow \infty)$  ここで  $C$  は一般に 0 でない。

即 (48) を (42) に代入すると

$$\delta'_{x',N} = \frac{2 \sum_{i=1}^N (2i-N) \left\{ \sum_{p=0}^k a_p \left( \frac{i}{n} \right)^p \right\}}{N^2}$$

二重級数の交換可能の条件の下で  $N \rightarrow \infty$  とすれば

$$\delta'_x = \lim_{N \rightarrow \infty} \delta_{x',N'} = \sum_{p=1}^k 2^{-p+1} a_p B(p+1, 2) \tag{50}$$

又均等線との交わりの回数、 $X=0$  及び  $X=1$  を除くと  $k-1$  回ありこれにより次数が分る。

次の (B) (C) にみるように周期性やランダム変動に対しては  $\delta_{x',N} \rightarrow 0$  であるから個数  $N$  が充分大なる時、(39) を充す (1) の系列曲線は明かに

$$M(X) = a_0 X + \frac{a_1}{2} X^2 + \cdots + \frac{a_{k+1}}{k+1} X^{k+1} \tag{51}$$

(B) (1) が周期性をもつ時

$$x_i = \sum_{i=1}^n (\alpha_i \cos 2\pi \lambda_i t + \beta_i \sin 2\pi \lambda_i t) \tag{52}$$

につき

$$\delta_{x',N} \sim 0 \left( \frac{1}{N} \right) \tag{53}$$

は明である。

(C) (1) が random variation の実現を示すならば iterated log の法則から

$$\delta_{x',N} \sim o(1) \tag{54}$$

を得る。

### § 3. 相関曲線と相関々係

相関々係を示すものとして相関曲線\*の概念を § 1, § 2 の方向に於て構成してみる。

(i) ベクトル系列

$$X_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \quad \dots, \quad X_N = \begin{pmatrix} x_N \\ y_N \end{pmatrix} \tag{55}$$

に対し

$$S_{x,n} = \sum_{i=1}^n x_i, \quad S_{y,n} = \sum_{i=1}^n y_i \tag{56}$$

を用い相関曲線を次のように定義する。

**定義 9**  $y$  の  $x$  に関する相関曲線  $K_{y,x,N}(X)$  とは

$$K_{y,x,N}(X) = \begin{cases} 0 & 0 \leq X < S_{x,1} \\ \frac{S_{y,n}}{S_{y,N}} & \frac{S_{x,n}}{S_{x,N}} \leq X < \frac{S_{x,n+1}}{S_{x,N}}; n = 1, 2, \dots, N-1 \\ 1 & X = 1 \end{cases} \tag{57}$$

と区間  $[0, 1]$  で定義される関数である。

即我々は第 13 図と同様の曲線を得るが、この場合、対角線との間の面積  $A$  は折線に対し、対角線の下の方々にある時を正、上方にある時を負とすれば、

$$\begin{aligned}
 2A_{x,y,N} &= \frac{\sum_{i=1}^n x_i \sum_{j<i} y_j - \sum_{i=1}^n y_i \sum_{j<i} x_j - \sum_{i=1}^n x_i y_i}{S_{x,N} \cdot S_{y,N}} \\
 &= \frac{\sum_{\substack{i>j \\ i,j=1,\dots,N}} (x_i y_j - y_i x_j)}{S_{x,N} \cdot S_{y,N}} = \frac{1}{N^2} \sum_{\substack{i>j \\ i,y=1,\dots,N}} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} \\
 &= \frac{1}{N^2} \sum_{\substack{i<j \\ i,j=1,2,\dots,N}} \overline{X_i \times X_j} = \frac{1}{N^2} \sum_{i,j=1,2,\dots,N} \text{sgn}(i, j) \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} \quad (58) \\
 &= \frac{1}{N^2} \sum_{\substack{i<j \\ i,j=1,2,\dots,N}} \overline{X_i \times X_j} = \frac{1}{N^2} \sum_{i,j=1,2,\dots,N} \text{sgn}(i, j) \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}
 \end{aligned}$$

である。

ここで  $X_i \times X_j$  は  $X_i, X_j$  のベクトル積又  $\overline{X_i \times X_j}$  をその積の大きさを表すことにする。

§ 2 と同様我々は

$$2A_{y,x,N} = D_{y,x,N} \quad (59)$$

$$\begin{cases} \frac{1}{N(N-1)} \sum_{i,j=1,2,\dots,N} \text{sgn}(i, j) \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} = \mathcal{D}_{y,x,N} \\ \frac{1}{N^2} \sum_{i,j=1,2,\dots,N} \text{sgn}(i, j) \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} = \mathcal{D}'_{y,x,N} \end{cases} \quad (60)$$

と定義するとき

$$-1 \leq D_{y,x,N} \leq 1$$

であり、 $D_{y,x,N}$  は系列  $y/x$  が順序づけられている程 1 又は  $-1$  に近く、不規則的である程 0 に近いからこれは比率の順位相関を支える。然しこの値は  $y/x$  の集中度に依存するから、 $y/x$  のジニ係数  $G_{y/x,(N)}$  を用いて

$$d_{y,x,N} = \frac{D_{y,x,N}}{G_{y/x,(N)}} = \frac{\mathcal{D}_{y,x,N}}{\mathcal{A}' y/x} \quad (61)$$

により標準化した方がよい。

なぜならば

$G_{y/x,(N)} \geq D_{y,x,N}$  従って  $0 \leq d_{y,x,N} \leq 1$  が成立するからである。

(ii) 今 vector  $(x, y)$  の各成分が区間  $[0, 1]$  でそれぞれ  $u$  の連続な正值関数として

$x = x_u, y = y_u$  と表わされる時

$x, y$  についての相関曲線は

$$\left\{ \begin{aligned} X &= \frac{\int_0^t x \, du}{\int_0^1 x \, du} \\ K(X) &= \frac{\int_0^t y \, du}{\int_0^1 y \, du} \end{aligned} \right. \quad (62)$$

と定義することが出来る。

この時

第 5 表 産業別資産の share distributions と集中係数 (Gini 係数)

産 業	推定法人数 *	推定資産評価額 * (百万円)	集中係数 G**	Pareto 常数 α***	$\frac{\sigma}{\mu}$
全 産 業	357272	6153141	0.88469	1.06517	24.53954
農 林 業	3523	4125	0.95898	1.02139	8.53403
水 産 業	1286	43656	0.81410	1.11418	10.78741
鉱 業	668	191581	0.93435	1.03513	4.98658
石 炭 業	75	123705	0.69710	1.21726	1.90098
そ の 他 鉱 学	593	67876	0.94300	1.03023	7.91028
建 設 業	16070	144384	0.85304	1.08614	23.19385
製 造 業	109007	2601331	0.90767	1.05086	18.04110
食 料 品 製 造 業	22560	230301	0.81454	1.11384	11.08739
紡 績 業	17110	467934	0.89680	1.05754	15.28596
木材及び木製品製造業	21079	144448	0.58626	1.35286	2.22943
紙及び類以品製造業	5614	142972	0.96360	1.01889	12.69969
化 学 工 業	8503	466723	0.92195	1.04233	10.41299
ガラス及び土石製品製造業	4012	93720	0.88867	1.06264	14.45083
第一次金属製造業	5306	544512	0.95622	1.02289	17.26955
金属製品製造業	7585	43118	0.85806	1.08271	10.74839
機 械 製 造 業	12982	541073	0.94949	1.02660	14.38368
その他の製造業	4256	16525	0.53224	1.43943	5.60697
卸売及び小売業	180737	1400452	0.72551	1.18918	8.25166
卸 売 業	83719	1077512	0.73984	1.17582	6.31216
小 売 業	97018	322940	0.55342	1.40348	13.16724
百 貨 店 業	161	51806	0.75974	1.15812	3.16998
その他の小売業	96857	271133	0.47058	1.56252	18.92804
金融および保険業	13544	170852	0.94813	1.02736	18.92804
銀行及び信託業	86	105917	0.69770	1.21664	2.07389
保 險 業	3564	10354	0.98660	1.00679	16.79095
その他の金融及び保険業	9894	54580	0.83233	1.10072	12.09409
不 動 産 業	4693	43466	0.93016	1.03854	9.88277
公 益 事 業	5768	1388739	0.97516	1.01274	13.12603
運 輸 業	5636	404523	0.93462	1.03497	8.71483
地方鉄道及び軌道業	106	138497	0.63586	1.28634	1.93758
道 路 運 送 業	2640	48904	0.78641	1.13580	3.01286
水 運 業	799	174580	0.94608	1.02850	5.59630
倉 庫 業	1008	17265	0.68422	1.23076	5.65814
その他の運輸業	1082	25275	0.90312	1.05363	13.94602
通 信 学	43	14358	0.75584	1.16151	2.57376
動 力 供 給 業	89	969857	0.85599	1.08412	2.44890
高 気 業	17	905199	0.39853	1.75460	0.80532
ガ ス 業	72	64657	0.91266	1.04785	4.62680
サ ー ビ ス 業	21976	164551	0.75077	1.16598	11.25720

\* data 源：中山一郎編 日本の国富構造

\*\* G は集中曲線の面積より算出した

\*\*\* α は  $\frac{1}{2G} + \frac{1}{2}$  によつて推定した

$$K'(X) = \frac{\int_0^1 x \, du}{\int_0^1 y \, du} \frac{y_t}{x_t} \quad (63)$$

であるから、 $K(X)$  は比  $y/x$  の系列曲線であり、従つて  $y = \lambda x$  のとき  $(0, 0)$  (11) を結ぶ対角線がこれに対応する。

又 (58) に対応して

$$\begin{aligned} D_{y,x} &= 2 A_{y,x} = \int_0^1 X dK(X) - \int_0^1 K(X) dX \\ &= \frac{1}{\mu_x \mu_y} \int_0^1 \left\{ \int_0^t x_u du \right\} y_t dt - \frac{1}{\mu_x \mu_y} \int_0^1 \left\{ \int_0^t y_u du \right\} x_t dt \\ &= \frac{1}{\mu_x \mu_y} \int_0^1 \int_0^1 \begin{vmatrix} x_u & x_t \\ y_u & y_t \end{vmatrix} du dt \end{aligned}$$

一方

$$\int_0^1 \left\{ \int_0^1 x_u du \right\} y_t dt = \int_0^1 \left\{ \int_0^1 y_u du \right\} x_t dt = \mu_x \mu_y$$

であるから

$$D_{y,x} = \frac{1}{\mu_x \mu_y} \int_0^1 \left\{ \int_t^1 y_u du \right\} x_t dt - \int_0^1 \left\{ \int_t^1 x_u du \right\} y_t dt = \frac{1}{\mu_x \mu_y} \int_0^1 \int_t^1 \begin{vmatrix} x_t & x_u \\ y_t & y_u \end{vmatrix} du dt$$

である。

従って結局

$$D_{y,x} = \frac{1}{2} \frac{1}{\mu_x \mu_y} \int_0^1 \int_0^1 \operatorname{sgn}(t, u) \begin{vmatrix} x_u & x_t \\ y_u & y_t \end{vmatrix} du dt \quad (64)$$

となる、ここで  $\operatorname{sgn}(t, u)$  は (47) に於けるものと同義である

又

$$|D_{x,y}| \leq \frac{1}{2} \int_0^1 \int_0^1 \left| \det \begin{pmatrix} x_u & x_t \\ y_u & y_t \end{pmatrix} \right| du dt$$

は明でありこれが  $x/y$  の Gini 係数を与える。又

$$\vartheta_{x,y} = \int_0^1 \int_0^1 \operatorname{sgn}(t, u) \begin{vmatrix} x_u & x_t \\ y_u & y_t \end{vmatrix} du dt \quad (64')$$

と定義出来る。

§ 4 多次元集中曲面と多次元系列, 相関曲面

価値 vector の系列又は集合に対し 1 次元の場合と同様集中曲面, 系列, 相関曲線等が定義可能である。

[I] 多次元集中曲面

(i) 非連続型の二次の vector 系列の出現頻度が表のように与えられた場合

第6表

	$x$	$x_1$	$x_2$	.....	$x_n$	
$y$	$y_1$	$f_{11}$	$f_{12}$	.....	$f_{1n}$	$f_{1\cdot}$
	$y_2$	$f_{21}$	$f_{22}$	.....	$f_{2n}$	$f_{2\cdot}$
	.....	.....	.....	.....	.....	.....
	$y_m$	$f_{m1}$	$f_{m2}$	.....	$f_{mn}$	$f_{m\cdot}$
		$f_{\cdot 1}$	$f_{\cdot 2}$	.....	$f_{\cdot n}$	

$$\begin{cases} X_{uv} = \sum_{i=1}^u \sum_{j=1}^v f_{ij} \\ Y_{uv} = \frac{1}{\mu_x} \sum_{i=1}^u \sum_{j=1}^v x_i f_{ij} \\ Z_{uv} = \frac{1}{\mu_y} \sum_{i=1}^u \sum_{j=1}^v y_j f_{ij} \end{cases} \quad (65)$$

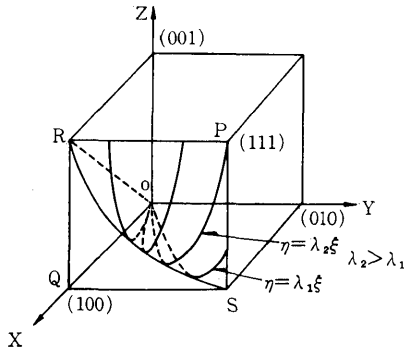
を結ぶ集中折面  $A(X, Y, Z) = 0$  が考えられる。

この時この面と角錐  $OPRQS$  が囲む体積のうち内側のものを  $V$  とすれば正観測値に対し

$$0 \leq 3V \leq 1$$

であり  
従って

$$G_{x,y} = 3V \quad (66)$$



第 14 図 二次元集中曲面

を二次元の  $G_{i_n}$  係数と定義出来る。

一方  $V = \frac{N(N-1)(N-2)A_{xy}}{18N^3\mu_x\mu_y}$  が証明出来る

ここで

$$A_{xy} = \frac{1}{3! \binom{N}{3}} \sum_{\substack{i \neq j, k \\ j \neq k \\ i, j, k=1, 2, \dots, N}} \left| \det \begin{pmatrix} x_i - x_k & x_j - x_k \\ y_i - y_k & y_j - y_k \end{pmatrix} \right| \quad (67)$$

でありこれは二次元の平均差と定義出来る [13].

以上は更に任意次元の vector  $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$  に対し拡張出来るが、その時は

$$X_0 = \sum_{1 \leq i_k \leq u_k} \sum_{1 \leq i_k \leq u_k} f_{i_1, i_2, \dots, i_n}$$

$$X_p = \frac{1}{\mu_{x_p}} \sum_{1 \leq i_k \leq u_k} \dots \sum_{1 \leq i_k \leq u_k} x_{i_p} f_{i_1, i_2, \dots, i_n} \quad (68)$$

$p=1, 2, \dots, n$

で表わされる  $n+1$  次元空間の点を結ぶ折面が集中折面  $A(X_0, X_1, \dots, X_n) = 0$  であり、平均差は  $N$  々の有限系列に対して

$$A_{x_1 x_2 \dots x_n} = \frac{1}{(n+1)! \binom{N}{n+1}} \sum_{\substack{i_k \neq i_h \text{ for } h \neq k \\ i_1, i_2, \dots, i_{n+1}=1, 2, \dots, N}} \sum \left| \det \begin{pmatrix} x_{i_1 1} - x_{i_{n+1} 1} & x_{i_1 2} - x_{i_{n+1} 2} & \dots & x_{i_1 n} - x_{i_{n+1} n} \\ x_{i_2 1} - x_{i_{n+1} 1} & x_{i_2 2} - x_{i_{n+1} 2} & \dots & x_{i_2 n} - x_{i_{n+1} n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i_n 1} - x_{i_{n+1} 1} & x_{i_n 2} - x_{i_{n+1} 2} & \dots & x_{i_n n} - x_{i_{n+1} n} \end{pmatrix} \right| \quad (69)$$

$$G_{x_1 x_2 \dots x_n} = (n+1)V = \frac{(n+1)! \binom{N}{n+1} A_{x_1 x_2 \dots x_n}}{N^{n+1} \prod_{i=1}^n \mu_{x_i}} \quad (69')$$

と与えられ、正值観測値に対し  $0 \leq G_{x_1 x_2 \dots x_n} \leq 1$  となる。

この形は S. S. Wilks [12] の multidimensional statistical scatter の形と対照的である。

(ii)  $x, y$  の連続な密度関数  $f(x, y)$  が存在するとき、

(65) に対応する  $(x, y)$  の集中曲面  $A(X, Y, Z) = 0$  は

$$\begin{cases} X = F(x, y) \\ Y = \frac{1}{\mu_x} \int_{-\infty}^x \int_{-\infty}^y \xi f(\xi, \eta) d\xi d\eta \\ Z = \frac{1}{\mu_y} \int_{-\infty}^x \int_{-\infty}^y \eta f(\xi, \eta) d\xi d\eta \end{cases} \quad (70)$$

によって定義出来る。

この時 (66) に相応する  $G_{xy}$  は

$$G_{xy} = \frac{A_{xy}}{6\mu_x\mu_y} \tag{71}$$

となる.

もし  $f(x, y)$  が独立な確率密度  $X, Y$  の密度関数であると考えられる場合には

$$\begin{aligned} A_{xy} &= 6 \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^x (x - \xi) f_1(\xi) d\xi \right\} f_1(x) F_1(x) dx \\ &\quad \times \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^y (y - \eta) f_2(\eta) d\eta \right\} f_2(y) F_2(y) dy \\ &= \frac{3}{2} \left\{ A_x - \int_{-\infty}^{\infty} A_{\xi| \xi \leq x} f_1(x) dx \right\} \left\{ A_y - \int_{-\infty}^{\infty} A_{\eta| \eta \leq y} f_2(y) dy \right\} < \frac{3}{2} A_x A_y \end{aligned}$$

が成立し又  $x, y$  の間に完全相関が認められる時

$$A_{xy} = 0$$

となる. 一般には右辺の各 factor を  $d_x, d_y$  で示すならば

$$0 \leq \frac{2}{3} \frac{A_{xy}}{d_x d_y} \leq 1$$

が証明出来る [13]

従って

$$C = 1 - \frac{2}{3} \frac{A_{xy}}{d_x d_y} \tag{73}$$

を新たな相関係数とすることが出来る.

一方 (67) に当る二次元 mean difference は

$$A_{x_1 x_2} = \int_0^1 \int_0^1 \int_0^1 \left| \det \begin{pmatrix} x_u - x_s & x_v - x_s \\ y_u - y_s & y_v - y_s \end{pmatrix} \right| dF_u dF_v dF_s \tag{74}$$

である.

以上をより一般的に  $n$  次元 vector  $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$  に適用するならば

$$\begin{aligned} X_0 &= F(x_1, x_2, \dots, x_n) \\ X_p &= \frac{1}{\mu_{xp}} \int \dots \int_{-\infty \leq \xi_k \leq x_k} \xi_p f(\xi_1, \xi_2, \dots, \xi_n) d\xi_1 d\xi_2 \dots d\xi_n \\ &\quad p=1, 2, \dots, n \end{aligned} \tag{75}$$

と表わした時に

$$A(X_0, X_1, \dots, X_n) = 0 \tag{76}$$

が集中曲面を与える.

且 この際 generalized mean difference は

$$\begin{aligned} A_{x_1, x_2, \dots, x_n} &= \int_0^1 \int_0^1 \dots \int_0^1 \left| \det \begin{pmatrix} x_{11} - x_{n+11} & x_{21} - x_{n+11} & \dots & x_{n1} - x_{n+11} \\ x_{12} - x_{n+12} & x_{22} - x_{n+12} & \dots & x_{n2} - x_{n+12} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} - x_{n+1n} & x_{2n} - x_{n+1n} & \dots & x_{nn} - x_{n+1n} \end{pmatrix} \right| \\ &\quad dF(x_{11}, x_{12}, \dots, x_{1n}) dF(x_{21}, x_{22}, \dots, x_{2n}) \dots dF(x_{n+11}, x_{n+12}, \dots, x_{n+1n}) \end{aligned} \tag{77}$$

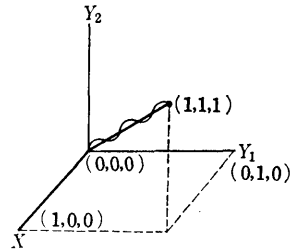
となる.

[II] 多次元系列曲面

(i) 集中曲線と同様系列曲線を多次元化することが出来る。

今二次元の vector 系列 (55) を考えた場合 (56) の記号を用いて

$$\begin{cases} X = \frac{\nu}{N} \\ Y = \frac{S_{x,\nu}}{S_{x,N}} \\ Z = \frac{S_{y,\nu}}{S_{y,N}} \end{cases} \quad \nu = 1, 2, \dots, N \quad (78)$$



第 15 図 二次元系列曲線

を結ぶ折線を

$$\begin{pmatrix} Y \\ Z \end{pmatrix} = M(X) = \begin{pmatrix} M_1(X) \\ M_2(X) \end{pmatrix} \quad (79)$$

で表わした時これをもって、二次元の系列折線と定義する。又この折線を含む折面が系列折える。そしてこの折面は常に  $x, y$  の二次元集中折面の上部に位置するのである。そして面をもし (78) を折線で示した場合はその projection  $Y = M_1(X)$ ,  $Z = M_2(X)$  はそれぞれ  $x, y$  に関する一次元系列折線を示し他の projection  $Z = K(Y)$  は一次元相関折線を示すことになる。

この際 (42) で与えた係数  $\delta_{x,N}$  は

$$\delta_{x,N} = \frac{1}{3! \binom{N}{3}} \sum_{\substack{i \neq j, k \\ j \neq k \\ i, j, k = 1, 2, \dots, N}} \text{sgn}(i, j, k) \begin{vmatrix} x_i - x_k & x_j - x_k \\ y_i - y_k & y_j - y_k \end{vmatrix} \quad (80)$$

$$\delta_{xy,N} = \frac{1}{N^3} \sum_{i,j,k=1,2,\dots,N} \text{sgn}(i, j, k) \begin{vmatrix} x_i - x_k & x_j - x_k \\ y_i - y_k & y_j - y_k \end{vmatrix} = \frac{3! \binom{N}{3}}{N^3} \delta_{x,N} \quad (80')$$

と一般化出来る。

ここで

$$\text{sgn}(i, j, k) \quad (81)$$

は  $i, j, k$  を大きさに順に並べ変える置換の偶奇に応じて  $+1$  又は  $-1$  の値をとるものとする。

$\delta_N$  に於じて  $C_N$  の一般化も生ずが、その際

$$-1 \leq C_{x,N} = \frac{\delta_{xy,N}}{\delta_{x,N}} \leq 1 \quad (82)$$

が成立する。

一次元の場合の  $C_{x,N}$  は  $x, y$  の順位相関係数の機能を持つものに対し、二次元の  $C_{x,N}$  は寧ろ相関係数の機能を有している。つまり  $y = \lambda x + \beta$  と表わされる時  $\delta_{x,N} = 0$  となるからである。

又相関曲線  $Z = K(Y)$  に関連して

$$\begin{aligned} \delta_{x,N} &= \frac{1}{3! \binom{N}{3}} \sum_{\substack{i \neq j, k \\ j \neq k \\ i, j, k = 1, 2, \dots, N}} \text{sgn}(i, j, k) \begin{vmatrix} x_i - x_k & x_j - x_k \\ y_i - y_k & y_j - y_k \end{vmatrix} \\ &= \frac{1}{\binom{N}{3}} \sum_{\substack{i < j < k \\ i, j, k = 1, 2, \dots, N-1}} \left\{ \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} - \begin{vmatrix} x_i & x_k \\ y_i & y_k \end{vmatrix} - \begin{vmatrix} x_k & x_j \\ y_k & y_j \end{vmatrix} \right\} \end{aligned}$$



$$= \frac{1}{\binom{N}{3}} \left\{ \sum_{i < j} \sum_{i=1,2,\dots,N-1} \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix} - \begin{vmatrix} x_{N-1} & x_N \\ y_{N-1} & y_N \end{vmatrix} - \sum_{i=1}^{N-2} \begin{vmatrix} x_i & x_N \\ y_i & y_N \end{vmatrix} \right. \\ \left. - \sum_{k \geq 3} \begin{vmatrix} x_1 & x_k \\ y_1 & y_k \end{vmatrix} + \sum_{j=2}^{N-2} \begin{vmatrix} x_j & x_{j+1} \\ y_j & y_{j+1} \end{vmatrix} + \begin{vmatrix} x_{N-1} & x_N \\ y_{N-1} & y_N \end{vmatrix} \right\}$$

こゝで (60) 式を用いると

$$\delta_{xy,N} = \vartheta_{y,x,N} + \frac{1}{\binom{N}{3}} \left\{ (x_N - x_1) N \mu_N - (y_N - y_1) N \mu_x + \sum_{i=1}^{N-1} \begin{vmatrix} x_i & x_{i+1} \\ y_i & y_{i+1} \end{vmatrix} \right\} \quad (83)$$

が成立する.

以上は亦  $n$  次元の vector 系列  $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$  に拡張出来るがその際は

$$\begin{cases} X_0 = \frac{\nu}{N} & \nu = 1, 2, \dots, N \\ X_i = \frac{S_{xi}}{S_{xi,N}} & i = 1, 2, \dots, n \end{cases} \quad (84)$$

を結ぶ折線を

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = M(X) = \begin{pmatrix} M_1(X) \\ M_2(X) \\ \vdots \\ M_n(X) \end{pmatrix} \quad (85)$$

で表わした時これをもって  $n$  次元系列折面と定義する. 又この折線を含む折面を系列折面とする.

この際同様に

$$\delta_{x_1 x_2 \dots x_n, N} = \frac{1}{(n+1)! \binom{N}{n+1}} \sum_{i_1, i_2, \dots, i_{n+1} = 1, 2, \dots, N} \sum_{i_k = i_k \text{ for } k \neq n} \dots \sum \begin{vmatrix} x_{i_{11}} - x_{i_{n+1}} & x_{i_{12}} - x_{i_{n+1}} & \dots & x_{i_{1n}} - x_{i_{n+1}} \\ x_{i_{22}} - x_{i_{n+1}} & x_{i_{23}} - x_{i_{n+1}} & \dots & x_{i_{2n}} - x_{i_{n+1}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{i_{1n}} - x_{i_{n+1}} & x_{i_{2n}} - x_{i_{n+1}} & \dots & x_{i_{nn}} - x_{i_{n+1}} \end{vmatrix} \quad (86)$$

$$\delta'_{x_1 x_2 \dots x_n, N} = \frac{(n+1)! \binom{N}{n+1}}{N^{n+1}} \delta_{x_1 x_2 \dots x_n, N} \quad (86')$$

が定義出来る.

(ii) (45) 及び (62) 式に準じて vector  $(x, y)$  の各成分が区間  $[0, 1]$  でそれぞれ  $u$  の連来な正值関数である場合二次元の系列曲線, 面を定義出来る. 即

$$X = t$$

$$Y = M_1(X) = \frac{\int_0^t x du}{\int_0^1 x du} \quad 0 \leq t \leq 1 \quad (87)$$

$$Z = M_2(X) = \frac{\int_0^t y \, du}{\int_0^1 y \, du}$$

とした時

$$\begin{pmatrix} Y \\ Z \end{pmatrix} = M(X) = \begin{pmatrix} M_0(X) \\ M_2(X) \end{pmatrix} \tag{88}$$

を  $x, y$  に関する二次元系列曲線と定義する. 又これを含む折面を系列折面とする.

この時 (80) (80') に当る係数  $\delta_{xy}$  は (47) の一般化として

$$\delta_{xy} = \int_0^1 \int_0^1 \int_0^1 \operatorname{sgn}(s, t, u) \begin{vmatrix} x_s - x_u & x_t - x_u \\ y_s - y_u & y_t - y_u \end{vmatrix} ds \, dt \, du \tag{89}$$

が妥当する. ここで  $\operatorname{sgn}(s, t, u)$  の定義は (81) で用いたのと同様である.

更に  $n$  次元の場合に拡張すれば

$$X_0 = t$$

$$X_i = M_i(X) = \begin{cases} \int_0^t x_{iu} \, du & 0 \leq t \leq 1 \\ \int_0^1 x_{iu} \, du & i = 1, 2, \dots, n \end{cases} \tag{90}$$

とした時に

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = M(X) = \begin{pmatrix} M_1(X) \\ M_2(X) \\ \vdots \\ M_n(X) \end{pmatrix} \tag{91}$$

によって  $n$  次元系列曲線を定義することが出来, 従って又曲面も定義出来る.

この時又

$$\delta_{x_1 x_2 \dots x_n} = \int_0^1 \int_0^1 \dots \int_0^1 \operatorname{sgn}(t_1, t_2, \dots, t_{n+1})$$

$$\times \begin{vmatrix} x_{t_1 1} - x_{t_{n+1} 1} & x_{t_2 1} - x_{t_{n+1} 1} & \dots & x_{t_n 1} - x_{t_{n+1} 1} \\ x_{t_1 2} - x_{t_{n+1} 2} & x_{t_2 2} - x_{t_{n+1} 2} & \dots & x_{t_n 2} - x_{t_{n+1} 2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t_1 n} - x_{t_{n+1} n} & x_{t_2 n} - x_{t_{n+1} n} & \dots & x_{t_n n} - x_{t_{n+1} n} \end{vmatrix} dt_1 \, dt_2 \dots dt_n \tag{92}$$

と定義することが出来る. その際

$$\operatorname{sgn}(t_1, t_2, \dots, t_{n+1}) \tag{93}$$

は  $t_1, t_2, \dots, t_{n+1}$  をその大きさに順に並べかえる置換の偶奇により  $+1$  又は  $-1$  をとるものとなる.

[III] 多次元相関曲面

(85) 乃至 (91) で支えられる vector 系列又は変数  $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$  は  $n+1$  次元空間  $\begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_n \end{pmatrix}$  の超

平面  $\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$  上に project される. それは又一般相関曲線従って又曲面を規定する.

(i)  $n$  次元 vector 系列  $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$  に対し (85) の定義が適用されると

$$K(X_1, X_2, \dots, X_n) = 0 \tag{94}$$

なる関係が成立するがこれを  $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$  の  $n$  次元相関曲線という。

この場合 (59) に於ける  $\vartheta_{xy,N}$  の一般化として

$$\vartheta_{x_1 x_2 \dots x_n, N} = \frac{1}{n! \binom{N}{n}} \sum_{t_1=t_2, t_3, \dots, t_n} \dots \sum_{t_i=t_j \text{ for } i=j} \text{sgn}(t_1 t_2 \dots t_n) \begin{vmatrix} x_{t_{11}} & x_{t_{21}} & \dots & x_{t_{n1}} \\ x_{t_{12}} & x_{t_{22}} & \dots & x_{t_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t_{1n}} & x_{t_{2n}} & \dots & x_{t_{nn}} \end{vmatrix} \quad (95)$$

$$\text{又 } \vartheta'_{x_1 x_2 \dots x_n, N} = \frac{n! \binom{N}{n}}{N^n} \vartheta_{x_1 x_2 \dots x_n, N} \quad (95')$$

こゝで  $\text{sgn}(t_1 t_2 \dots t_n)$  は (93) 式に準ずるものとする。

(ii) もし (91) で与えられた系列曲線に対し (94) の関係が得られる場合には同様にして

$$\vartheta_{x_1 x_2 \dots x_n} = \int_0^1 \int_0^1 \dots \int_0^1 \text{sgn}(t_1, t_2, \dots, t_n) \begin{vmatrix} x_{t_{11}} & x_{t_{21}} & \dots & x_{t_{n1}} \\ x_{t_{12}} & x_{t_{22}} & \dots & x_{t_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t_{1n}} & x_{t_{2n}} & \dots & x_{t_{nn}} \end{vmatrix} dt_1 dt_2 \dots dt_n \quad (96)$$

これが (64') の generalization である。

### III 集団の作用

#### § 1. 静態的作用

What are the forces governing business concentration and how have these changed is recent years? [6]

この古典的の調題に対しては競争システムの相互作用についての理解を必要とする。

我々はこれを競争力等同一もしくは grade 極めて近いでの作用と寡占, 独占力等 grade 間の作用に分けて考える。

(i) grade 内の作用

微分可能な集中曲線  $A(x)$  のある grade 区間  $[a, b]$  における平均差  $\Delta(a, b)$  及び集中係数  $G(a, b)$  はこの部分の集中曲線  $A(x; a, b)$  が

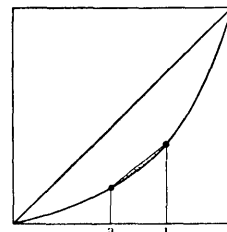
$$A(x; a, b) = \frac{A\{(b-a)X + a\} - A(a)}{A(b) - A(a)} \quad (97)$$

$$0 \leq x \leq 1$$

である事から容易に次式で示すことが出来る。

$$G(a, b) = 1 - 2 \frac{\int_a^b \{A(X) - A(a)\} dX}{(b-a)\{A(b) - A(a)\}} \quad (98)$$

$$\Delta(a, b) = 2 \frac{\mu(a, b)}{\mu} G(a, b) = 2 \frac{A(b) - A(a)}{b-a} G(a, b) \quad (99)$$



第 16 図 truncate された集中曲線

従って集中曲線の  $X$  grade の近傍  $[X, X + \Delta X]$  での平均差, 集中係数を求めるには, 上式 (98), (99) で  $a = X, b = X + \Delta X$  とすればよい。事実独占度 degree of monopoly では  $G(0.9, 1)$  又は  $\Delta X$  として上位数社等が用いられているのである。

処で P. E Hart (I) によれば平均差  $\Delta$  及び集中度係数は集団における degree of dominance を表わす数量であった。このことから一歩進めてこの逆数をもって degree of competition

を、又微少区間  $[X, X + \Delta X]$  に対して  $\lim_{\Delta X \rightarrow 0} \frac{G(X, X + \Delta X)}{\Delta X}$  の逆数をもつ grade  $X$  に於ける競争力  $f_w$  を定義出来ないだろうか。ただし一般に競争は grade が近いもの程完全に作用すると考えられているからである。

この時 (97) (98) (99) の関係で

$$\left. \begin{aligned} g(X) &= \lim_{\Delta X \rightarrow 0} \frac{G(X, X + \Delta X)}{\Delta X} = \frac{A'(X)}{6A'(X)} \\ d(X) &= \lim_{\Delta X \rightarrow 0} \frac{\Delta(X, X + \Delta X)}{\Delta X} = \frac{A'(X)}{3} \end{aligned} \right\} \quad (100)$$

を得るから、上記の考えにより  $f_w$  は

$$c(X) = \frac{1}{g(X)} \quad (101)$$

に比例する。

(ii) grade 間の作用

寡古的競争のシステムでは一般に grade 間又システムの grade に対する作用が存在する。それは grade 内の作用が grade によって異なる原因でもある。

一方で分布法則の成立はその内部に何等かの静態的均衡の存在が前提にあるのではなからうか。equilibrium の認識は Gini の基本的認識でもあると受取れる。従つてもし  $X$  が grade 間の作用  $f_b$  を受けるならば

$$f_w + f_b = 0 \quad (102)$$

が予想される。

もし

(A)  $f_b = 0$  ならば (100) 式の関係から一点分布。又は均等線が得られる。これは完全競争システムといえる。

(B)  $f_b = c$  (一定) ならば、同様に Zipf 分布  $f(x) = k/x; 0 < x$  を得る。

(C)  $f_b$  が grade に比例する、即

$$f_b = C_1 X + C_2$$

ならば一般の Pareto 型分布を得るが、パレート分布は

$$\frac{dc(X)}{dX} * = -6a \quad (103)$$

を充すから  $\frac{dC(X)}{dX} = \lambda \frac{df_w}{dX} = -\lambda \frac{df_b}{dX} = -\lambda C_1$  を考えると

$$c_1 = 6\lambda a \text{ 従つて } f_b = 6\lambda a X + c_2$$

を得る。従つて Pareto 常数  $a$  は、この system における一種の重力恒数といえ、競争力の変動の要因となる。

(D) もし (103) に於て競争力の変化  $c'(X)$  がその grade の企業規模によ比例するならば即

$$c'(X) = C_1 A'(X) + C_2$$

ならばこの分布は  $\Gamma$  分布となり

(E) 規模の対数に比例するとき Gibrat 分布を得る。

## § 2. 動態的作用 [14]

(i) 集中現象を齎す数量は一般に諸要因の多次元的構成をもち (其等要因自体が集中現象従つて grade 内及び grade 間の作用が認められるが) その構成様式によって次の時点の現象形態を決定する。

ある二数量  $x, y$  の間に一定の弾力性 ( $y$  の  $x$  に関するその係数を  $\beta$  とする) が認められ

\*  $\frac{dc(X)}{dX}$  は密度分布で表わすと  $\frac{d \log f(x)}{d \log x}$  に等しい。[10]

るとき、それは (100) で定義された  $c(X)$  を用いると

$$\beta = \frac{c_y(X)}{c_x(X)} \tag{104}$$

で表わされる。つまり恒数  $\beta$  の存在は競争力、従って又 grade 間の作用がこの二量に関して grade に拘りなく一定の比率を持って作用している状態、従ってこのシステムの  $x, y$  に関する一種の安定性、即  $x, y$  に関する均衡状態を示すことになる。

扱具体的に今企業資産の集中曲線の時間的変動を考えることにする。

資本の集中、集積は、生産力の発展に依存し、逆に生産力の発展段階は資本の蓄積規模に依存することが一般に容認されているから、今問題とするシステム内の各企業に Cobb-Douglas 型の生産関数

$$P = b L^k K^{1-k} \tag{105}$$

が成立していると前提する。

ここで  $P$  は生産物の数量、 $L$  は労働用役量  $K$  は固定資本量とする。又  $k$  は常数であつて従って  $P$  と  $L$  又は  $K$  との間に一定の弾力性があると仮定されている。

資本集約度  $\gamma = K/L$  を用いると上式は更に

$$P = b K \frac{1}{\gamma^{1-k}} \tag{106}$$

で示される。

最も単純な場合としてこの産出量  $P_t$  が次の時点の固定資本量  $K_{t+1}$  と一定弾力性係数  $\delta = \frac{E K_{t+1}}{E P_t}$  もって結ばれているとすれば

$$K_{t+1}^{1/\delta} = a K_t \gamma^{k-1}$$

であり更にデータによるとこの技術進歩度或は生産力段階を示す  $\gamma$  は又  $K_t^\varepsilon$  に比例している [第 17 図参照] つまり一定の弾力性をもつことから

$$K_{t+1} = c K_t^{(1+\varepsilon)(k-1)\delta} \tag{107}$$

となる。

即  $K_{t+1}$  の分布は II. § 1, (III) の例に示したように、 $K_t$  の分布が Pareto 係数  $\alpha$  をもつ Parets 型であるとすれば

$$f(K_{t+1}) = \lambda \frac{f(K_t)}{K_t^{(1+(k-1)\varepsilon)\delta-1}} = \mu \frac{1}{K_t^{\alpha_t(1+(k-1)\varepsilon)\delta}}$$

従って

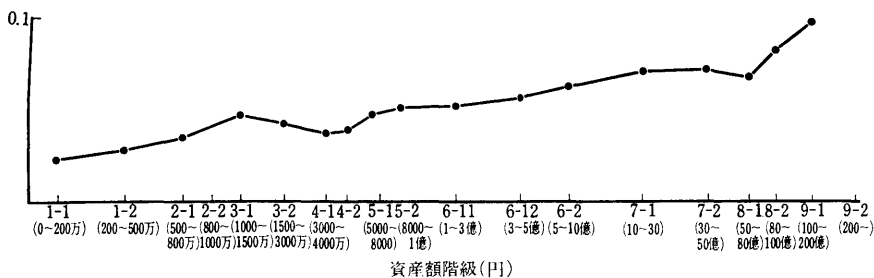
$$\alpha_{t+1} = \alpha_t + \{1 + (k-1)\varepsilon\}\delta - 1 \tag{108}$$

を与える。

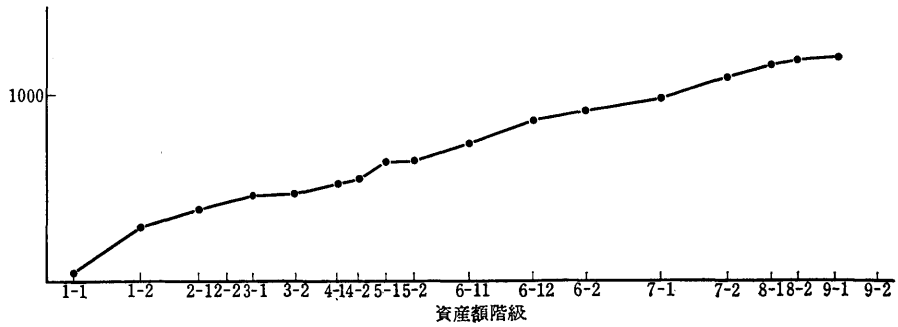
一般に  $da/dt$  は以上の関数モデルによれば労働生産性、資本化率、資本集約度等に依存し、その関係を反映しているのである。

従って又  $\alpha$  自体が集中段階と生産力の発展段階、そしてその過程を表わしている。

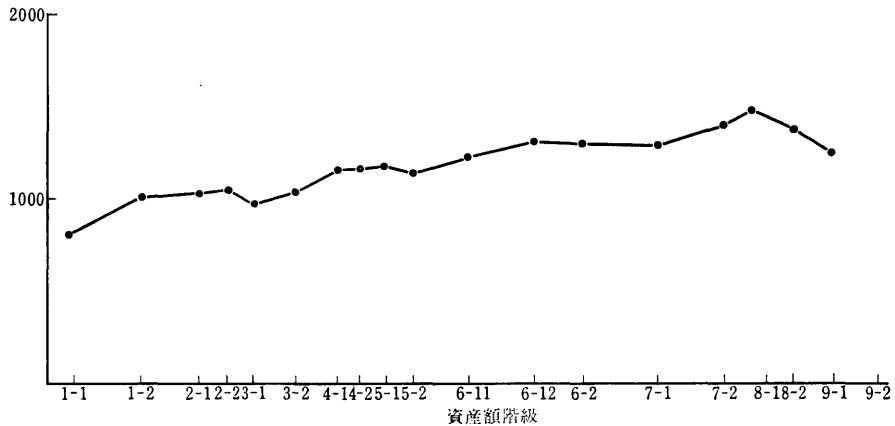
資本系数 (総資本) (i) 製造業総括 (資本額階級別)



資本集約度 (総資本) (ii) 製造業総括 (資産階級別)



労働生産性 (iii) 製造業総括 (資産階級別)

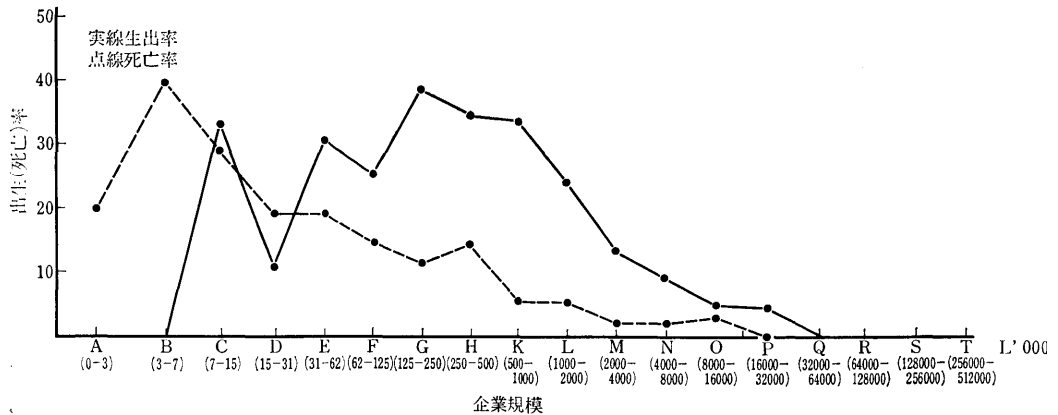


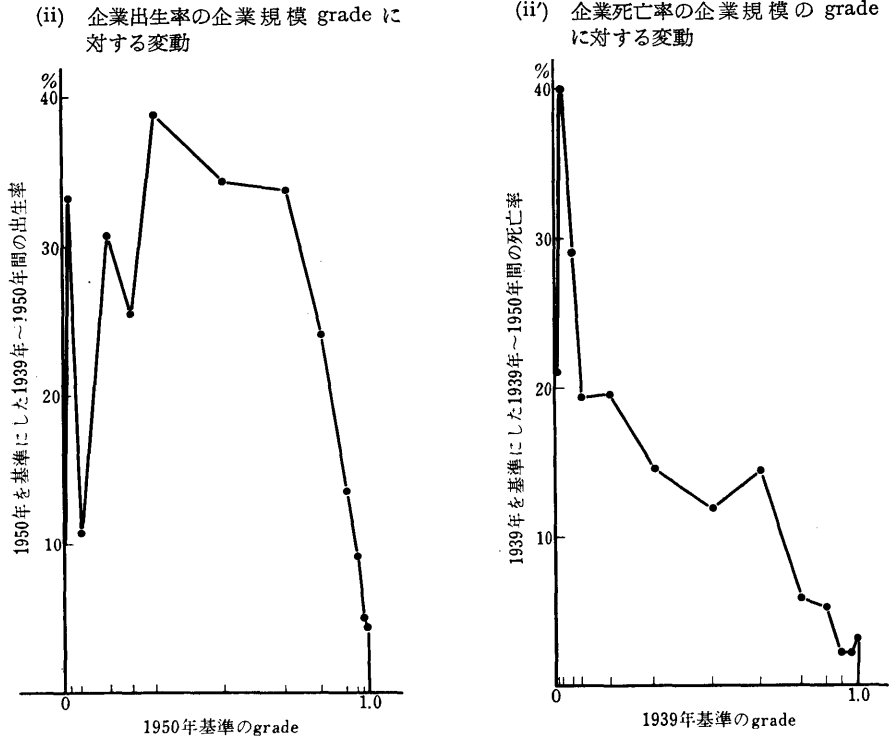
第 17 図 企規模別にみた諸係数の変動 (昭和 30 年度国審調査再集計結果: 経済企画庁経済研究, 資本構造と企業間格差, より作製)

勿論 Gibrst 分布が妥当するシステムに於ては Gibrat 係数が Pareto 係数に代つてその機能を果している。

(iii) 分布或は分配法則の成立は  $f_w$  と  $f_b$  の balance に求めることが出来たが、一方 grade に不確定性を齎す本質要因は前時点の grade 内作用即競争力と考えられる。即、競争力  $c(X)$  が大なる程行動における個的決定性を失うと同時に集团的決定性をますと思われるからである。

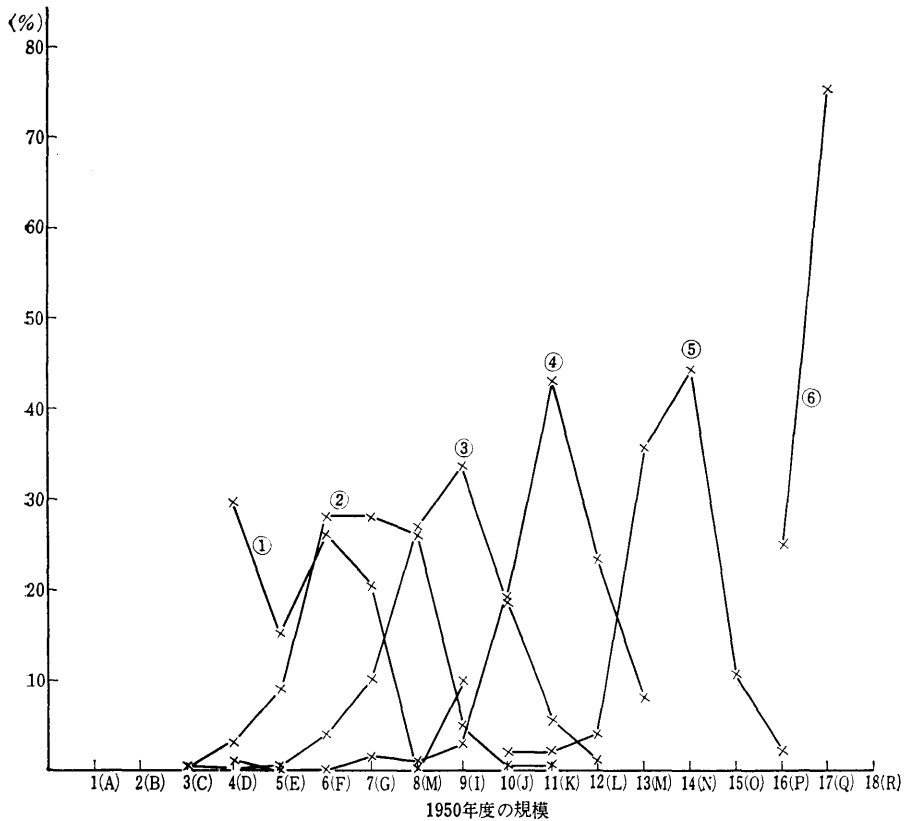
(i) 企業出生率及び死亡年の規模に対する変動 (1939 年~1950 年における英国の資料 [6] より算出)





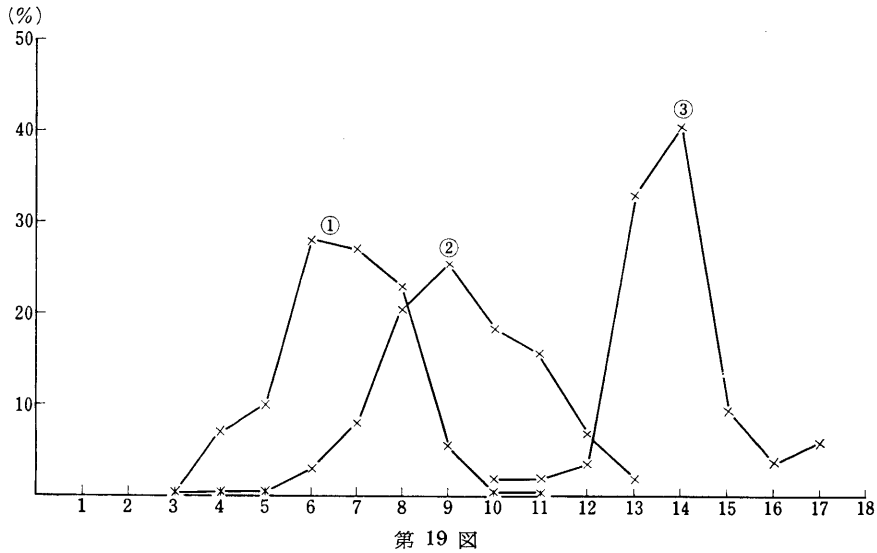
第 18 図

(i) 企業規模の推移 (1939 年を基準にした 1950 年における英国企業の階級推移—資料:[6]より算出)



分布の数字は 1936 年度のスケールを相対的に 3 階級ずつ繰り上げたもの

(ii) (i) において基準時の階層を 6 ケづつ一括した場合



例えば第 18 図の如く企業の出生、死滅のその grade の企業数に対する比率は、競争力に比例している、即不安定の度合は grade が上昇する程減少するからである。(因みにこの data 例では母集団は Gibrat 型に近い) このことは Xgrade 近傍内の企業は全体として  $f_b$  と balance を保つが個々の行動は統計的にしか、決定されないという見解によって、支持されるであろう。規模に比例する確率という概念はこうした事情の反映と受取れる。

#### IV 技術的諸方法

##### § 1 統計的推論への適用

###### [I] 標本平均差の分布

I, § 1 で与えた平均差  $A_N$  について Nair, U. S, (1936)[8] は連続な分布  $F(x)$  をもつ母集団からの  $n$  ケの標本による標本平均差  $A$  の分布を与えた。即  $A$  は順序統計量  $X_{(i)}$ ;  $i=1, 2, \dots, n$  を用いて

$$\begin{cases} A = \frac{2}{n(n-1)} \{2U_n - (n+1)V\} \\ \text{ここで } U = \sum_{j=1}^n \{j X_{(j)}\}, U = \sum_{j=1}^n X_{(j)} \end{cases} \quad (109)$$

と表わせるから

$$E(A) = 2 \int_{-\infty}^{\infty} x(2F-1) dF \quad (110)$$

$$\text{又 } E(A^2) = \frac{4}{n(n-1)} \{I_1 + 2(n-2)I_2\} \quad (111)$$

ここで

$$I_1 = \int_{-\infty}^{\infty} x^2 \{ (n-1) - 4(n-2)F + 4(n-2)F^2 \} dF$$

$$I_2 = \int_{-\infty}^{\infty} x^2 dx_2 \int_{-\infty}^{x_2} x_1 dF_1 \{ (n-3) - 2(n-5)F_1 - 2(n-1)F_2 + 4(n-3)F_1F_2 \}$$

である。但し、 $\alpha < 2$  をもつパレート分布 (34) では  $E(A^2)$  が発散することが分る。



然し

(i) 正規母集団

$$dF = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx; \quad -\infty \leq x \leq \infty$$

に対しては

$$E(\Delta) = \frac{2\sigma}{\sqrt{\pi}}$$

$$\text{Var.}(\Delta) = \frac{4\sigma^2}{n(n-1)} \left\{ \frac{n+1}{3} + \frac{2(n-2)\sqrt{3}}{\pi} - \frac{2(2n-3)}{\pi} \right\} \\ \sim \frac{\sigma^2}{n} (0.8068)^2$$

であり,

指数分布に従う集団

$$dF = \frac{1}{\tau} e^{-\frac{x}{\tau}} dx; \quad 0 \leq x \leq \infty$$

に対しては  $E(\Delta_1) = \tau$

$$\text{Var}(\Delta_1) = \tau^2 \frac{2(2n-1)}{3n(n-1)} \sim \frac{4}{3n} \tau^2$$

又矩形分布

$$dF = \frac{1}{k} dx \quad 0 \leq x \leq k$$

に従う母集団に対しては

$$E(\Delta_1) = \frac{1}{3} k$$

$$\text{Var}(\Delta_1) = \frac{k^2}{9} \frac{n+3}{5n(n-1)} \sim \frac{k^2}{45n}$$

等が計算されている。

[II] 平均差による Tchebycheff 型不等式  
Glasser, G. J [5] は

$$d' = \frac{1}{N^2} \sum_i \sum_j |x_i - x_j| \quad (i, j = 1, 2, \dots, N)$$

により, 母平均  $\mu$  及び  $d'$  が与られた時

$$P(\mu - k_1 d' < X < \mu + k_2 d') \geq 0 \quad (\text{但し } 2k_1 k_2 \leq k_1 + k_2 \text{ の時}) \\ \geq \sqrt{4k_1 k_2 - 2(k_1 + k_2)} / (k_1 + k_2) \\ (\text{但し } 2k_2(k_1 - k_2) \leq k_1 + k_2 \leq 2k_1 k_2 \text{ の時}) \\ \geq (2k_2 - 1) / 2k_2 \quad (\text{但し } k_1 + k_2 \leq 2k_2(k_1 - k_2) \text{ の時}) \\ (112)$$

を得た.

特に  $k_1 = k_2$  とすれば

$$\begin{cases} P(\mu - k d' < X < \mu + k d') \geq 0 & k \geq 1 \\ \leq \sqrt{(k-1)/k} & k \geq 1 \end{cases} \quad (113)$$

となる.

この他  $\Delta, G$  等については多くの関係が存在する.

例えば Iyengar, N. S [7] は Gibrat 分布

$$f(x; \theta, \lambda) = \frac{1}{x\lambda\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\log x - \theta}{\lambda} \right)^2} dx, \quad 0 < x < \infty$$

に対し (勿論此の場合  $E(X) = e^{\theta + \frac{1}{2}\lambda^2}$ ,  $\text{Var.}(X) = e^{2\theta + \lambda^2} (e^{\lambda^2} - 1)$  であるが) 集中曲線は  $\varphi(x)$  を規準正規分布とした場合, (9) 式の  $X, Y$  は

$$X = \varphi\left(\frac{\log x - \theta}{\lambda}\right) \quad Y = \varphi\left(\frac{\log x - \theta}{\lambda} - \lambda\right)$$

であり Gini 係数は

$$G = 1 - 2\varphi\left(-\frac{1}{\sqrt{2}}\right) = 2\varphi\left(\frac{\lambda}{\sqrt{2}}\right) - 1$$

であることから  $G$  の最大推定重  $\tilde{G}$  は  $\lambda$  のそれを  $\tilde{\lambda}$  とすれば

$$\tilde{G} = 2\varphi\left(\frac{\tilde{\lambda}}{\sqrt{2}}\right) - 1$$

を充すことを示した.

又  $\tilde{G}$  は漸近的に

$$E(\tilde{G}) = G, \text{Var.}(\tilde{G}) = \frac{\lambda^2 e^{-\frac{1}{2}\lambda^2}}{2\pi n}$$

をもつ正規分布に従うという結果を得た.

又 Muniruzzaman, A. N. M\* は Pareto 母集団からのランダムサンプルにより Pareto 恒数  $\alpha$  の最大推定量を求め, それが標本幾何平均  $G$ , 標本最小値  $X$  を用いると

$$\hat{\alpha} = \frac{1}{\log \frac{G}{X}}$$

であり, これは一貫性をもち又

$$\text{Var}(\hat{\alpha}) \rightarrow \frac{\alpha^2}{N-3}$$

であると報告している.

### § 2. サンプリング・メソッドへの適用

[I] 単調母集団におけるスタート・ランダムな系統抽出について

系統抽出については現に Madow [9] による系列相関係数を用いた計算が存在する.

即母集団 size を  $N$ , sample size を  $n$ , 抽出間隔を  $k$  とすれば

$N = kn$  であり, そのサンプルの一組を

$$S_i = \{x_i, x_{i+k}, \dots, x_{i+(n-1)k}\}$$

とした時

$$n \bar{x}_i = x_i + x_{i+k} + \dots + x_{i+(n-1)k} \tag{114}$$

と表すことが出来る.

$\bar{x}$  を標本平均とすれば  $S_i$  が Sample に当たった時

$$\bar{x} = \bar{x}_i \text{ であり}$$

この時系列相関係数を用いて

$$\left. \begin{aligned} \varepsilon \bar{x} &= \bar{x} \\ \rho_{\bar{x}}^2 &= \frac{\sigma^2}{n} \{1 + 2 \sum_{\mu} \rho_{k\mu}\} \\ &= \frac{\sigma^2}{n} \{1 + (n-1) \bar{\rho}_k\} \end{aligned} \right\} \tag{115}$$

が得られる.

然し, 歪み型母集団を対象とすれば既述の諸観点から以上の結果の適用は不能となる.

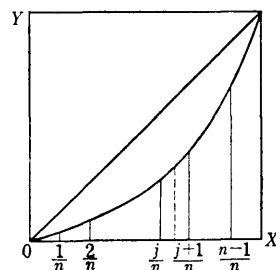
\* Calcutta Statist. Ass. Bull., 7 (1956-7)

今単調母集団の集中曲線  $A(X)$  を考えると、上記の等間隔抽出は  $X$  を  $1/n$  間隔で分割し各組からその下端 (或は上端) より  $i/N$  順位のものを抽出することになる。

母集団の極限分布が与えられその密度関数が微分可能と仮定出来る時、各組の中央値での  $A(X)$  の Taylor 展開を考えると母平均は

$$\begin{aligned} \mu &= \mu \sum_{j=0}^{n-1} \left\{ A\left(\frac{j-1}{n}\right) - A\left(\frac{j}{n}\right) \right\} \\ &= \mu \sum_{j=1}^{n-1} \left\{ \frac{1}{n} A'\left(\frac{j+\frac{1}{2}}{n}\right) \right. \\ &\quad \left. + \frac{2\left(\frac{1}{2n}\right)^3}{3!} A''\left(\frac{j+\frac{1}{2}}{n}\right) + \left(\frac{1}{n^5}\right) \right\} \quad (116) \end{aligned}$$

となり、この第1項は II に述べた如く  $A'(X) = x/\mu$  であるから  $\frac{\sum x_j}{n}$  となる。



第 20 図 系統抽出法

従って各組の中央値をサンプルとすれば、標本的平均  $x = \frac{\sum x_j}{n}$  を母平均の推定量とすることが出来る

その時、第二項は誤差項となり近似的に

$$\begin{aligned} &\sum_{j=1}^{N-1} \frac{1}{24 n^3} ((x_{i+1} + x_{i-1} - 2x_i)) \\ &= \frac{1}{24 n} (x_N + x_{N+1} - x_2 + x_1) \leq \frac{R}{12 n} \quad (R \text{ は range}) \quad (117) \end{aligned}$$

で評価出来るよう。

スタート・ランダムの際はその標本としたときのそれに対する bias は  $\frac{\sum_{i=0}^n (x_{i+1} - x_i)}{n} = \frac{R}{n}$  を越さないから、これを (117) に加えて推定値の評価が出来る

### [II] 非等確率抽出法

物価指数の算出等で RG. D. Allen [1] は家計支出費目の規模に比例した確率での抽出法を示唆しているが、これは支出費目の集中曲線  $Y = A(X)$  の  $Y$  軸の  $[0, 1]$  区間を等確率抽出しそれに該当する  $X$  で示される費目を対象とすることになる。この  $x$  を推定する際の補助変量は、勿論相関の高いものがよい。

又単純な分布関数の変換を想起すると、対数規模に比例した抽出法は  $\log X$  を等確率に抽出することになり、標本幾何平均、対数分散等を算出すれば母集団の幾何平均を推定出来る。例えば Gibrat 分布は Gauss 分布、Pareto 分布は指数分布の性質から推定出来る。

然し算術平均の推定は困難であり、それ故もし規模に比例した等間隔抽出法を用いる場合は [I] の展開を  $A^{-1}(X)$  について考えれば標本調和平均によるべきであることが分る。

### [III] 層別法

改良されたジニの平均差

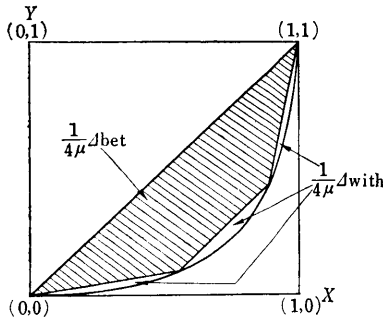
$$D' = \frac{1}{N^2} \sum_{i,j=1}^n |x_i - x_j|$$

は分散と同じくこれを層内分散  $D'$  with と層間分散  $D'$  bet に分解出来る。

$$\begin{cases} D' = D' \text{ with.} + D' \text{ bet.} \\ D' \text{ with.} = \sum_{i=0}^n p_i^2 D_i' \quad D_i' \text{ は } i \text{ 層の平均差 } p_i \text{ は } i \text{ 層の密度 } \frac{N_i}{N} \\ D' \text{ bet.} = \sum_{i,j=1}^n p_h p_k |\mu_i - \mu_j| \end{cases} \quad (118)$$

である。

つまり  $N$  個の有限系列の集中折線について第 20 図の分解を考えれば上式の結果が得られる。



第 21 図 平均差の分解

一般に連続な密度関数の集中曲線の truncation の性質 ([10] 及び第 20 図参照) として例えば grade  $\xi$  で上下に二分したとき、下部の集団による集中曲線及び平均を  $A_1(x; \xi)$ ,  $\mu_1(\xi)$ , 上部のそれを  $A_2(x; \xi)$ ,  $\mu_2(\xi)$  とすれば

$$\begin{cases} A(\xi) A_1(X; \xi) = A(\xi X) & 0 \leq X \leq 1 \\ \frac{\mu_1(\xi)}{\mu} = \frac{A(\xi)}{\xi} = A'(\xi X_\mu) \\ G_1(\xi) = 1 - 2 \frac{\int_0^\xi A(X) dX}{\xi A(\xi)} \end{cases} \quad (119)$$

又

$$\begin{cases} \{1 - A(\xi)\} A_2(X; \xi) + A(\xi) = A\{(1 - \xi) X + \xi\} & 0 \leq X \leq 1 \\ \frac{\mu_2(\xi)}{\mu} = \frac{1 - A(\xi)}{1 - \xi} = A'\{\xi + (1 - \xi) X\} \\ G_2(\xi) = 1 - 2 \frac{\int_\xi^1 \{A(X) - A(\xi)\} dX}{(1 - \xi)\{1 - A(\xi)\}} \end{cases} \quad (120)$$

及び

$$A_{bet} = \frac{1}{2} \{\xi - A(\xi)\} \quad (121)$$

となるから  $\frac{d A_{bet}}{d \xi} = 0$  をとけば最適な層の分割として、規模順に隣接した二層を加えた平均値が層の境界になっていけばよいことが分る。

具体的に上例では

$$A'(\xi) = 1$$

即全体の平均値そのものが境界であり、一般に  $n$  層に分つ時は、その境界を  $\xi_0 = 0, \xi_1,$

$\xi_2, \dots, \xi_n = 1$  とすれば  $\frac{\partial A_{bet}}{\partial \xi_i} = 0, i = 1, 2, \dots, n - 1$  をといて

$$\frac{A(\xi_{i+1}) - A(\xi_{i-1})}{\xi_{i+1} - \xi_{i-2}} = A'(\xi_i) \quad i = 1, 2, \dots, n \quad (122)$$

を満足していればよいことが分る。

## 結 び

以上は過去数年に亘って日本統計学会、統計数理研究所等における報告及び本年度当研究年所附属養成所における体験に基づいて日本統計学会会報、当研究所彙報、Annals 等における論文及び養成所基本科テキストを製正加事した上、更に一層の理論的整備を加えたものである。

個々の結論は既に屢触れたので省略するが、全般的にみて此の派の方法は形式的には非確率論的、又非正規型の統計といいきれるかもしれない。然し例えば序文に触れた O. R 部門に關聯するというならば、両者は対象に於て接近するが、此の部門の統計が著しく工学的実用的であり短期的部分的方法である。此の派の統計はその本質的展開に於てはいはばより基礎的であり、長期的全体的観点に立つものであり、実証的方向にあるということを本文の展開の過程に於て感じ取れるのである。此は故 Gini 教授の関心の方向（それは主として人口、経済、生物学の諸領域にあったが）と態度に照してみても肯づける処であろう。 統計数理研究所

以上におけるグラフ作製，集計に関しては中島道子嬢の手を煩はした。深く感謝する。

## 参 考 文 献

- [1] R. G. D. Allen, "Sampling for current economic statistics", *J. R. S. S.*, series A, Vol. 127, 1964.
- [2] R. C. Blitz and J. A. Brittain, "An extension of the Lorenz Diagram to the correlation of two variable", *Metron*, Vol. XXIII, N. 1-4, 1964.
- [2] V. Castellano, "Corrado Gini: a memoir", *Metron*, Vol. XXIV N. 1-4, 1965.
- [4] C. Gini, "On the characteristics of Italian statistics", *J. R. S. S.*, Vol. 128, 1965.
- [5] G. J. Glasser, "Tchebycheff-type inequalities in terms of the mean difference". *Metron* Vol. XXIII, N. 1-4, 1964.
- [6] P. E. Hart and S. J. Prais, "The analysis of business concentration" *J. R. S. S.*, Vol. 119, 1956.
- [7] N. S. Iyengar, "On the standard error of the Lorenz concentration ratio", *Sankhya*, Vol. 22, 1960.
- [8] M. G. Kendall, *The Advanced Theory of Statistics*, Vol. 1, Charles Griffin & Co. Ltd., 1948.
- [9] L. M. Madow and V. G. Madow "On the theory of systematic sampling, I", *Ann. Math. Stat.* Vol. 15, 1944.
- [10] T. Taguchi, "Concentration-curve methods and structures of skew populations", *Ann. Ins. Stat. Math.*, Vol. 20, No. 1, 1968.
- [11] T. Taguchi, "On some properties of concentration curve and its applications," (to be applications", *Metron.*, Vol. XXV-N. 3-4, 1967.
- [12] S. S. Wilks, "Multidimensional statistical scatter", Olkin and others, *Contributions to Probability and Statistics*, Stanford Univ. Press, California, 1960.
- [13] 田口時夫, 多次元集中曲線, 平均差, ジニ係数 (予定)
- [14] H. S. Houthakker, "The Pareto distribution and the Cobb-Douglas production function in active analysis", *Review of Economic Studies*, 1956.