

二変量パレート分布の集中解析

—構造統計量の提案—

統計数理研究所 田 口 時 夫

(1982年12月 受付)

1. 序 論
 2. 集中解析の概念と方法
 3. 構造統計量の諸表現
 4. パレート集中曲面の構造統計量
 5. 結 語
- 付録：諸公式
参考文献

1. 序 論

パレート分布は、最初の所得分布として古典的価値ある分布であるが、解析的にみても、一次モメントを除いて、一般に高次モメントは存在しない点に於て従来の統計分布概念とそれに基づく解析方法論に対して衝撃的な役割を果たしたといえる (Pareto [17] 及び Johnson-Kotz [9] 参照). その反面此の分布は、ローレンツ曲線法、ジーニ集中係数に適切な表現と理解を見出すことに於て、イタリー学派の形成の機縁を与えた事は、此れ又周知の事実であろう (Lorenz [11], Gini [7] 参照). 然し乍らパレート分布の多変量化の試みは遙かに遅れて K.V. Mardia [14] によって 1962 年に至り漸く実現されることになった. 更にその後における多変量パレート分布に関する考察は、筆者の若干の報告を除くと、1980 年における B.C. Arnold による報告 [2] を知るのみである. 他方、ローレンツ曲線の多次元化の試みとしての集中曲面の形成は、筆者 [18], [19] と G. Lunetta [12], [13] の報告を除いては、若干の未完成と思われる報告を見るに留まる. 従って本報告は以上の二つの別個の方式の多変量化の試みであると共に、其等を再結合することによりどのように関係づけられるかを知る為の解析的実験であるが、その際に従来よく知られている母数とは異質な一群の母数 (筆者は文中でこれを構造母数及び構造統計量と名づけた) を析出し、その性格を解明し、その利用・計算法を呈示し得たことを併記せねばならない. このことに関連して以下の説明を加える必要が感じられる. すなわち筆者は 1982 年度日本数学会秋季大会及び 1982 年度科研費総合 A シンポジウムに於て、確率論に基礎づけられる大量観察としての統計的方法に対して、相互作用を併う集団現象を対象とする統計方法は、独自の基礎をもつべきであるとして、集団的構造及び集中構造を測度論的に規定した. 本稿で構造統計量といわれるものは、此の立場より直接演繹されるので、このような名称を与えたのである (田口 [23] 参照). 集中構造によって規定されるローレンツ曲線は、特定条件の下で密度分布によって定義される従来のローレンツ曲線の定式に還元されるのであるが、本稿では、従来通りの定式より出発することにして、基礎についての論及を避けることにした. つまりこのことは、ローレンツ曲線は確率密度とは無関係にその本質に基づいて拡張されるべきであるという筆者の観点を放棄するものではなく、此の種の見解を充分基礎づける為には、別個に一論文を必要とするからに他ならない. ここでは単に次の諸定義を提案するにとどめたい.

定義 1 「集合 X における σ 加法族 α に属する任意の集合 A に対して測度 $\mu(A)$ 及びベクトル値集合関数 $D(A)$ が与えられるものとする. この時 $\mu(A)$ 及び $D(A)$ の各成分 $D_i(A)$ が完全加法的である時 (X, α, μ, D) を集团的構造と定義する.*

定義 2 「今 X を R_n , 又は $D(A)$ が n 次ベクトル関数である集团的構造に於て, 特に $\mu(X) = 1, D_i(X) = 1, i = 1, 2, \dots, n$ 又 $(\mu(A), D_i(A))$ が各 i について平面上の非凸集合をなすとき, これを集中構造と定義する。」

X を企業集団の全体, A は特定産業, 又は特定規模階層, D は資産, 従業員数等の評価額と対応させて理解すれば $\mu(A)$ は規模分布を, 又 $D(A)$ はシェア構造に対応するものといえよう.

定義 3 「集中構造に於て, α の部分集合 \mathfrak{B} を次の B の全体とする. すなわち B を interval $I_i = (-\infty, x_i]$ 又は $I_i^c = [x_i, \infty)$ によって規定される集合とする. この時 $(\mu(B), D(B))$ によって構成される曲面を n 次集中曲面, 又その曲面に囲まれる体積を n 次集中度と定義する。」

1 次の集中構造に於て, 上記の集中曲面は曲線として複数の x の関数 $(\mu(x), D(x))$ によって表現される.

この時 $\mu(x), D(x)$ を微分可能であるとし,

$$(1.1) \quad D'(x) = d(x) \quad \mu'(x) = m(x)$$

とし, 且つ

$$(1.2) \quad \frac{d(x)}{m(x)} = w(x)$$

とすれば, 上記のローレンツ曲線は

$$(1.3) \quad f(y) = \frac{cm\{w^{-1}(cx)\}}{w'\{w^{-1}(cx)\}}$$

を用いて次節 (1) のケンドール式の経験的形式で表現されうる (田口 [23], [24] 参照). 集中構造が確定的構造といえる為には, n 次集中度が 0 であれば充分である. 逆に不確定的構造といえる為には n 次集中度が 0 でないことが必要である.

2. 集中解析の概念と方法

集中曲面とその性質及び利用については, 既に Taguchi [18] 及び Taguchi [22] に詳細に報告されているが, 第 4 節の解析の意義を明確にする為にここで再び其等の概観を与えることにする.

(1) 集中曲面の概念

一般に用いられる表現に従うと

(a) 例えば $0 < |E(X)| < \infty$ のとき, 連続な確率変数 X に対して定義されるローレンツ曲線はベクトル表現を用いると

$$(2.1) \quad L(x) = P(X \leq x) \left(1, \frac{E(X|X \leq x)}{E(X)} \right)$$

である. 従って

(b) $0 < |E(X)|, |E(Y)| < \infty$ のとき連続な確率ベクトル (X, Y) に対して定義されるロー

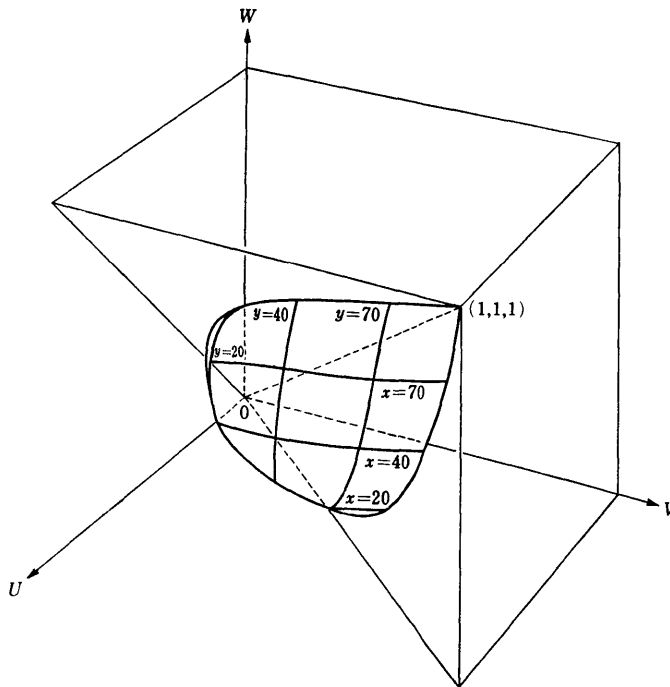
* いわゆる「統計的構造」は, 確率の集合を考えるものであり, ここで規定した集団なる構造とは本質的に異なる概念である.

二変量パレート分布の集中解析

レンツ曲面は

$$(2.2) \quad L(x, y) = P(X \leq x, Y \leq y) \cdot \left(1, \frac{E(X|X \leq x, Y \leq y)}{E(X)}, \frac{E(Y|X \leq x, Y \leq y)}{E(Y)} \right)$$

とすることが出来る。



第 2.1 図 集中曲面の一例

密度関数 $f(x, y) = \frac{1}{3.24x^{1.5}y^{1.5}}$; $1 \leq x, y \leq 100$ のパレート型分布の集中曲面

(2.2) 式の定義の結果第 2.1 表に抽象化された統計表は (2.2) 式に定義されたローレンツ曲面 (集中曲面) の一つの境界曲線に対応するものといえる。この事は此の曲面が実務統計表に密着した表現であり、その解析に一定の役割を果しうることを示すものといえる*。

第 2.1 表 実務統計表の理論的一表現

階級 項目	階 級 区 分						計
	$x < x_1$	$x_1 \leq x < x_2$	$x_{i-1} < x \leq x_i$	$x_{n-1} < x \leq x_n$	
構 成 比	u_1	u_2	u_i	u_{n-1}	1.0
x の分配比率	v_1	v_2	v_i	v_{n-1}	1.0
y の分配比率	w_1	w_2	w_i	w_{n-1}	1.0

* 「真に科学としての統計学は、従来の統計学がやめていたところではじまらなければならない」 ヴィットシュタイン (足利末男 [1] pp.366~367)

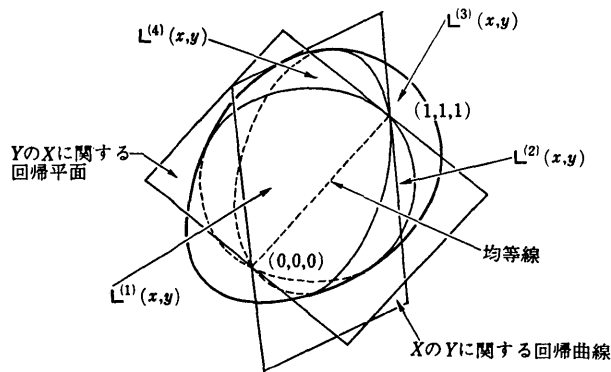
(2.2) 式の曲面を特に不完全又は狭義集中曲面とすることにする。それは(2.2)式に於て X, Y の x, y に対する不等関係を変えると(2.2)の類似曲面が得られるからである。(2.2)式を含めて此等を $L^{(1)}(x, y), L^{(2)}(x, y), L^{(3)}(x, y), L^{(4)}(x, y)$ とし、此等を併せると境界曲線を除き重複点のない閉じた曲面が得られる。これを完全集中曲面又は集中多様体としよう。具体的に今

$$L^{(1)}: X \leq x, Y \leq y \quad L^{(2)}: X \geq x, Y \leq y$$

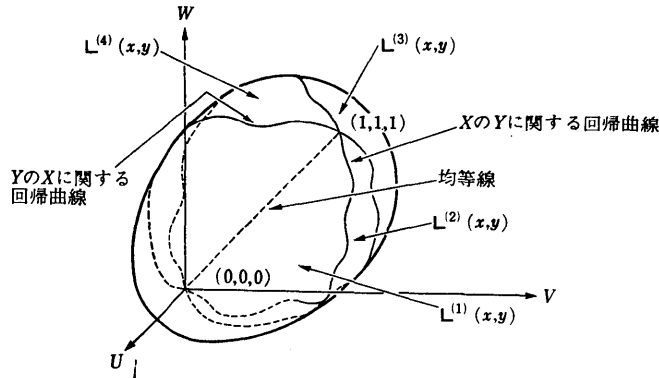
$$L^{(3)}: X \geq x, Y \geq y \quad L^{(4)}: X \leq x, Y \geq y$$

とすると、下図のような曲面となる。

i) Y の X 及 X の Y に関する回帰が exact linear の場合



ii) 非線型回帰をもつ一般の場合



(c) (b) に於て $L(x, \infty)$ 及び $L(\infty, y)$ を考えるとこれは曲面の境界曲線を与える。(2.2)式によると此等はそれぞれ Y の X 及び X の Y に関する回帰曲線の集中面上の対応曲線である。更に $\frac{\partial L(x, \infty)}{\partial x}$ を考えると、これは上記の理論的統計表に対応していることが明らかとなる。又 $L(x, y)$ を (U, V, W) 座標で考えると、 $L(x, \infty)$ の UV 座標面及び $L(\infty, y)$ の WU 座標面への正射影はそれぞれ X 及び Y のローレンツ曲線を与える。これに対して $L(x, \infty)$ の WU 座標面及び $L(\infty, y)$ の UV 座標面への正射影は、相関曲線 (Blitz and

Brittain [3]) といわれる曲線である。他方、此等の曲線の VW 面への正射影は、後述するように、比率曲線といえる性質をもっている。更に此等の正射影曲線が図に示す均等線の同一平面への正射影とで囲む面積は、それぞれジニ係数及びジニ様係数の $1/2$ を与えることになる。此等の係数をベクトルとして表現し、次項で示されるジニ・ベクトル $G_{y,x}, G_{x,y}$ とすることが出来る。 $G_{y,x}, G_{x,y}$ は、線形集中回帰係数及び集中相関係数を決定する (次項参照)。特に Y の X に関する回帰が線形ならば、 $L(x, \infty)$ は均等線を含む平面上にある。従ってこれを特に集中回帰平面ということが出来る。

(2) 集中曲面の性質

(1) の性質により、モメント相関・回帰に対応する諸関係を与えることが出来る。つまり

(a) (2.2) 式に於て特に

$$(2.3) \quad l(x, y) = \frac{\partial^2 L(x, y)}{\partial x \partial y}$$

とする。この時上述のジニ・ベクトルは、 $G_{y,x}$ に関しては

$$(2.4) \quad 2 G_{y,x} = \int_{(X_1, Y_1)} \int_{(X_2, Y_2)} [\text{sgn}(X_2 - X_1)] \cdot l(X_1, Y_1) \times l(X_2, Y_2)$$

と表わされる。 $G_{y,x}$ の第3成分は、 X のジニ係数となり、第2成分は Y の X に関する相関面積の2倍となる。又第1成分は X, Y 間の関係比率と X との相関を表わすことが後に指摘される。以上 (2.3) (2.4) を用いると Y の X に関する完全集中回帰直線 \tilde{y} は

$$(2.5) \quad G_{y,x} \cdot l(x, \tilde{y}) = 0$$

ここで

$$l(x, \tilde{y}) = \begin{pmatrix} F(x) \\ \frac{1}{E(X)} \int_{-\infty}^x \xi f(\xi) d\xi \\ \frac{1}{E(Y)} \int_{-\infty}^x \tilde{y} f(\xi) d\xi \end{pmatrix}$$

で定義される (Taguchi [22] 参照)。更に

$$(2.6) \quad G_{y,x} = (G_{y,x}^{(1)}, G_{y,x}^{(2)}, G_{y,x}^{(3)}) \\ G_{x,y} = (G_{x,y}^{(1)}, G_{x,y}^{(2)}, G_{x,y}^{(3)})$$

とすれば、 Y の X に関する、 X の Y に関する及び XY 間の集中相関係数はそれぞれ

$$(2.7) \quad \rho_{y,x} = \frac{G_{y,x}^{(2)}}{G_{y,x}^{(3)}} \quad \rho_{x,y} = \frac{G_{x,y}^{(2)}}{G_{x,y}^{(3)}} \\ \rho_{xy} = \pm \sqrt{\rho_{y,x} \rho_{x,y}} ; \\ \text{sgn } \rho_{x,y} = \text{sgn}(\rho_{x,y} + \rho_{y,x})$$

である (Taguchi [22] 参照)。

ところで、 Y の X に関する回帰が線形であるとき、すなわち $E(Y|X=x) = \beta x + \alpha$ ならば

$$(2.8) \quad \beta = \frac{E(Y) G_{y,x}^{(2)}}{E(X) G_{y,x}^{(3)}} \quad \alpha = \frac{E(Y) G_{y,x}^{(1)}}{G_{y,x}^{(3)}}$$

となる。従って(2.5)は Y の X に関する最小二乗回帰に一致する。更に $E(X|Y=y) = \beta'y + \alpha'$ も成立するならば、(2.7)式の ρ_{xy} は又モメント相関係数に等しい。然し乍ら(2.5), (2.7)は一次の絶対モメントの存在のみを前提とするから、モメント相関、回帰よりも広い適用性をもつことが理解出来る。

(b) (2.4), (2.5) を更に分析することにする。今、曲面(2.2)の一点 (x, y) における面積素分は

$$(2.9) \quad N(x, y) dx dy = \frac{\partial L(x, y)}{\partial x} \times \frac{\partial L(x, y)}{\partial y} dx dy \\ = \int_{-\infty}^x \int_{-\infty}^y l(x, \xi) \times l(\xi, y) d\xi d\zeta dx dy$$

である事を考えると(2.4)式の $G_{y,x}$ は微少表面積を大きさとする集中多様体の法線ベクトルの符号を顧慮した平均を示している。特に Y の X に関する回帰が exact に線形ならば $G_{y,x}$ は $L(x, \infty)$ を含む回帰平面の法線ベクトルを、又 $\|G_{y,x}\|$ は此の面上で $L(x, \infty)$ が囲む面積の大きさを与える。

(c) 完全集中曲面の囲む体積 V は

$$(2.10) \quad V = E_{(X_1, Y_1)} E_{(X_2, Y_2)} E_{(X_3, Y_3)} \{[\text{sgn}(X_1 - X_3) \text{sgn}(Y_2 - Y_3)] \cdot \\ \{l(X_1, Y_1) \times l(X_2, Y_2)\} \cdot l(X_3, Y_3)\} \div 4E(X)E(Y)$$

である。単調な回帰関係に於ては(この時 $\rho_{y,x} = 1$) $V = 0$ は残差がない状態を示す。従って V をもって二次元集中係数 G_{xy} とすることができる。つまり $X, Y \geq 0$ のとき

$$(2.11) \quad 0 \leq G_{xy} \leq 1$$

となる。以上により集中体積は Y の X に関する関係の不確定性の測度といえる。一方(2.5)の回帰残差 $R_{y,x}$ の一つの表現は

$$(2.12) \quad R_{y,x} = \frac{E|G_{y,x} l(X, Y)|}{G_x}$$

である。(2.4), (2.10), (2.11)の形をみれば $R_{y,x}$ は一つの体積量であると理解される。残差 $R_{y,x}$ の他種の表現は

$$(2.13) \quad R_{y,x} = E_{(X, Y)} (\|\pm N^{(i)}(x, y) - G(y, x)\|)$$

符号は $i=1, 2$ のとき正, $i=3, 4$ のとき負である。

(3) 構造統計量の概念と性質

前節(2.7), (2.8), (2.10)式等で与えられた諸量の構成を更に分析することにしよう。この場合(2.6)式の $G_{y,x}$ の各成分が容易に次式で示されることに注目する必要がある。すなわち $\mu_x = E(X)$, $\mu_y = E(Y)$ とした時、

$$(2.14) \quad G_{y,x}^{(1)} = \frac{A_{x,y,x}}{2\mu_x \mu_y} \\ G_{y,x}^{(2)} = \frac{A_{y,x}}{2\mu_y} \\ G_{y,x}^{(3)} = \frac{A_x}{2\mu_x} = G_x$$

ここで $(X_1, Y_1), (X_2, Y_2)$ は凡て同一分布に従うものとして

$$(2.15) \quad \begin{aligned} A_{x,y,x} &= \frac{E}{(X_1, Y_1)} \frac{E}{(X_2, Y_2)} \left[\{\text{sgn}(X_1 - X_2)\} \begin{vmatrix} X_1 & X_2 \\ Y_1 & Y_2 \end{vmatrix} \right] \\ A_{y,x} &= \frac{E}{(X_1, Y_1)} \frac{E}{(X_2, Y_2)} [\{\text{sgn}(X_1 - X_2)\} (Y_1 - Y_2)] \\ A_x &= \frac{E}{X_1} \frac{E}{X_2} [|X_1 - X_2|] \end{aligned}$$

である。同様の表現は $G_{x,y}$ についても可能である。従って (2.7), (2.8) 式から容易に

$$(2.16) \quad \begin{aligned} \rho_{y,x} &= \frac{A_{y,x}}{A_y}, & \rho_{x,y} &= \frac{A_{x,y}}{A_x} \\ \rho_{xy;\Delta} &= \pm \sqrt{\frac{A_{y,x} A_{x,y}}{A_x A_y}}, & \text{sgn } \rho_{xy;\Delta} &= \text{sgn}(\rho_{x,y} + \rho_{y,x}) \end{aligned}$$

$$(2.17) \quad \beta = \frac{A_{y,x}}{A_x}, \quad \alpha = \frac{A_{x,y,x}}{A_x}$$

を得ることが出来る。

他方 (2.10) 式は Taguchi [18] により

$$(2.18) \quad G_{xy} = V = \frac{A_{xy}}{4\mu_x\mu_y}$$

$$(2.19) \quad A_{xy} = \frac{E}{(X_1, Y_1)} \frac{E}{(X_2, Y_2)} \frac{E}{(X_3, Y_3)} \left[\{\text{sgn}(X_1 - X_3) \text{sgn}(Y_2 - Y_3)\} \begin{vmatrix} 1 & 1 & 1 \\ X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \end{vmatrix} \right]$$

と表される。

ここで $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ は凡て同一分布に従うものである。以上の結果によると (2.15) 式及び (2.19) 式に示される平均差 A_x 及びその類似諸形式がモメントに類似した役割を果たすものであることが理解されよう。従って我々はここで $A_x, A_{y,x}, A_{x,y,x}, A_{xy}$ 等の一群の諸形式を一括して構造母数としてこれまでの諸統計量と識別することを提言するのが妥当と考えるのである。又この時 (2.14), (2.16), (2.17), (2.18) の各式に示された諸量は構造統計量と呼称されて然るべきであろう。実例として一変量の正規分布に対しては

$$(2.20) \quad \sigma_x = \frac{\sqrt{\pi}}{2} G_x$$

であることがよく知られており、又容易に二変量の正規分布に対して

$$(2.21) \quad V = \frac{3(1 - \rho_{xy}^2)}{2\sqrt{4 - \rho_{xy}^2}} G_x G_y \quad (\text{Taguchi [22] 参照})$$

が得られる。ここで ρ_{xy} はモメント相関であると共に (2.16) 式に示される集中相関 $\rho_{xy;\Delta}$ でもあることは既述の通りである。つまりこの場合モメント統計諸量に対して、構造統計諸量が具体的且つ簡明に対応づけられているのである。

更にモメントに直接的対応をもたない構造母数 $A_{y,x,x}$ に特別に注目する必要がある。此の母数を理解する為、今見方を変えて比率 $\frac{Y}{X}$ を考察することにしよう*1。註に示すように、統

*1 「統計数学の利用者は、すべて個々の数字にたいして意味がないことを知っている。他の統計量と比較し関係づけることによって、はじめて統計数字は生命を得るのである」 §14 比率の節冒頭 p.142 (Flaskämper [6] 参照)。

計的比率に関して一般的比率は個別比率の分母による加重平均である*2。このような分母による加重期待値を E' によって表現することにする。この時密度関数 $f(x)$ が連続ならば

$$(2.22) \quad E'\left(\frac{Y}{X}\right) = \int_{-\infty}^{\infty} \frac{y}{x} \frac{x}{E(X)} f(x) dx \\ = \frac{E(Y)}{E(X)}$$

であり、一般的比率を与えていることが理解出来る。

よって更にこの加重期待値概念をもとにして得られる平均差等を A', ρ' で表現することになると、(2.16)に対応して

$$(2.23) \quad \rho_{y/x, x}' = \frac{A'_{y/x, x}}{A'_{y/x}}$$

となる。ここで

$$(2.24) \quad A_{y/x, x}' = \frac{E'}{(X_1, Y_1)} \frac{E'}{(X_2, Y_2)} \left[\operatorname{sgn}(X_1 - X_2) \left(\frac{Y_1}{X_1} - \frac{Y_2}{X_2} \right) \right] \\ = \frac{E}{(X_1, Y_1)} \frac{E}{(X_2, Y_2)} \left[\operatorname{sgn}(X_1 - X_2) \begin{vmatrix} Y_1 & Y_2 \\ X_1 & X_2 \end{vmatrix} \right] \\ = -A_{x, y, x}$$

同様にして

$$(2.25) \quad A_{y/x}' = \frac{E}{(X_1, X_2)} \frac{E}{(Y_1, Y_2)} \left(\begin{vmatrix} Y_1 & Y_2 \\ X_1 & X_2 \end{vmatrix} \right)$$

である。

従って $A_{x, y, x}$ は (2.17) 式を通じて回帰直線を決定する一方、加重比率平均差を示す (2.25) 式を分母とし、負の符号をつけることによって比率 $\frac{Y}{X}$ とその分母との加重相関係数を与えるのである。境界曲線 $L(x, \infty)$ の VW 座標面への正射影と考えられる曲線が比率に一定の関係をもつことは、既に米沢氏 [25] 等が指摘されている処である。

このような相関係数は以下に述べるように極めて實際上意義あるものと思われるので $\rho_{y/x, x}'$ を特に比率係数 $r_{y, x}$ と呼ぶことを提案したい。つまりまず歴史的に振り返ってみると、エンゲル係数が法則的地位を勝ち得たのは、特にこの係数が分母である所得に対して負の相関をもつことを見出したからといえないであろうか (Engel [5] 参照)。又例えば今日離職率や雇用率が分母となる従業員数に対して正又は負の相関をもつことが示されるとしたならば、これに大きな社会的関心を寄せざるを得なくなるであろう。同様な例は賃金率・持家比率又施設の稼働率、遊休比率等に多くの例をひくことが出来よう。逆に比率係数を用いて、有効な統計的比率を見出したり、又統計的比率の有効性を比較する基準とすることができるであろう。因みに、二変量対数正規分布

$$(2.26) \quad f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\right. \\ \left. \times \left\{ \frac{(\log x - \mu_x)^2}{\sigma_x^2} - \frac{2\rho(\log x - \mu_x)(\log y - \mu_y)}{\sigma_x\sigma_y} + \frac{(\log y - \mu_y)^2}{\sigma_y^2} \right\} \right]$$

*2 「個々の個別比率は一般比率よりも一部が大きく一部が小さいということは先にみた。厳密に言えば一般比率は個別比率の平均、特に加重平均である。」 p.173 (Flaskämper [6] 参照)

Φ を標準正規分布関数としたとき、次の結果を得ることができる。

$$(2.27) \quad G_x = \frac{A_x}{2\mu_x} = 2\Phi\left(\frac{\sigma_x}{\sqrt{2}}\right) - 1$$

$$(2.28) \quad G_{y,x}^{(2)} = \frac{A_{y,x}}{2\mu_y} = 2\Phi\left(\frac{\rho\sigma_y}{\sqrt{2}}\right) - 1$$

$$(2.29) \quad G_{y,x}^{(1)} = \frac{A_{x,y,x}}{2\mu_x\mu_y} = -2\Phi\left(\frac{(\beta_{y,x}-1)\sigma_x}{\sqrt{2}}\right) + 1$$

ここに $\beta_{y,x}$ は y の x に関する対数回帰係数を示す。更に

$$(2.30) \quad G_{y/x'} = \frac{A_{y/x'}}{2E'\left(\frac{Y}{X}\right)} = 2\Phi\left(\frac{\sqrt{\sigma_x^2 - 2\rho\sigma_x\sigma_y + \sigma_y^2}}{\sqrt{2}}\right) - 1$$

従って、 y の x に関する比率係数を $r_{y,x}$ とすると

$$(2.31) \quad -1 \leq r_{y,x} \leq 1$$

となり、 $\rho = \pm 1$ のとき、つまり $\frac{y}{x}$ と x が単調な関係をもつとき ± 1 となる (田口 [24] 参照)。

又、一方

$$(2.32) \quad \beta_{y,x} \geq 1$$

に対して

$$(2.33) \quad r_{y,x} \geq 0$$

となる。 $\beta_{y,x} = 1$ は $\frac{y}{x}$ が一定である条件である。

最後に、 y の x に関する集中相関係数 $\rho_{y,x}$ 、比率係数 $r_{y,x}$ を用いると (2.5) 式の集中回帰は (2.16), (2.17), (2.25) 式により

$$(2.34) \quad \hat{y} = \rho_{y,x} \frac{A_y}{A_x} x + r_{y,x} \frac{A_{y/x'}}{A_x}$$

と表わされる*3。

3. 構造統計量の諸表現

構造母数、構造統計量が集団分析上、極めて意味深い内容をもつことは、前節に述べた通りであるが、此等の統計量は何れも二重及至六重の期待値によって表現される点で、一般的に解析上の困難が予想されるであろう。たしかにモメント及びモメント統計量に較べて此等の計算は遙かに複雑になる。然し乍ら一変数の集中解析の場合と同様にまたもや「一見したところ非常に大きな計算上の労力はかなり簡単にされる。」(Flaskämper [6] p. 140) ののである。次節のような集中解析を容易にする母数の表現形式は、対象とする分布によって異なる選択が必要である。いわゆるジーン平均差にあたる一変量に関する二次元 (面積) 平均差 A_* は周知のように

*3 「ただ二つの例だけをあげるならば、相関の計算と傾向の計算はもちろん批判的に用いられるならば、社会統計学者に閉ざされている知識を与えることが出来る。しかしこの際数理統計学の方法を社会統計学の武器として取り上げるだけでなく、その事物論理的意味を研究することは、ドイツ統計学、否、統計学一般におわされた基本的課題である」 p. 337 (Flaskämper [6] 参照)

$$\begin{aligned}
 (3.1) \quad d_x &= \mathbf{E}_{X_1, X_2} (|X_1 - X_2|) \\
 &= \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 > X_2) - X_2\} P(X_1 > X_2)] \\
 &\quad - \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 < X_2) - X_2\} P(X_1 < X_2)] \\
 &= \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 > X_2) - X_2\} P(X_1 > X_2)] \\
 &\quad - \mathbf{E}_{X_2} [\{\mathbf{E}(X) - \mathbf{E}(X_1 | X_1 > X_2) P(X_1 > X_2)\} - \{\mathbf{E}(X) - X_2 P(X_1 > X_2)\}]
 \end{aligned}$$

であるから結局

$$(3.2) \quad d_x = 2 \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 > X_2) - X_2\} P(X_1 > X_2)]$$

同様に

$$(3.3) \quad d_x = -2 \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 < X_2) - X_2\} P(X_1 < X_2)]$$

従って又 (3.2), (3.3) より

$$\begin{aligned}
 (3.4) \quad d_x &= \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 > X_2) - X_2\} P(X_1 > X_2)] \\
 &\quad - \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1} (X_1 | X_1 < X_2) - X_2\} P(X_1 < X_2)]
 \end{aligned}$$

となり、三つの表現形式 (3.2), (3.3), (3.4) が得られる。此等に対応した表現形式を以下それぞれ第1, 第2, 第3形式としよう。

(1) 二変量に関する二次元 (面積) 平均差

この場合は前節にのべた d_x の類似平均差 $d_{y,x}, d_{x,y,x}$ は、上記の d_x と同様な方法によりそれぞれ次の第1~第3形式をもつ。

$$\begin{aligned}
 (3.5) \quad d_{y,x} &= \mathbf{E}_{(X_1, Y_1), (X_2, Y_2)} [\{\text{sgn}(X_1 - X_2)\} (Y_1 - Y_2)] \quad (\text{原形式}) \\
 &= \mathbf{E}_{X_1, Y_1} [\{\text{sgn}(X_1 - X_2)\} \{\mathbf{E}_{Y_2} (Y_2 | X_1) - \mathbf{E}_{Y_2} (Y_2 | X_2)\}] \\
 &= 2 \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1, Y_1} (\mathbf{E}_{Y_2} (Y_2 | X_1 > X_2) - \mathbf{E}_{Y_2} (Y_2 | X_2)) P(X_1 > X_2)\}] \quad (\text{第1形式}) \\
 &= -2 \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1, Y_1} (\mathbf{E}_{Y_2} (Y_2 | X_1 < X_2) - \mathbf{E}_{Y_2} (Y_2 | X_2)) P(X_1 < X_2)\}] \quad (\text{第2形式}) \\
 &= \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1, Y_1} (\mathbf{E}_{Y_2} (Y_2 | X_1 > X_2) - \mathbf{E}_{Y_2} (Y_2 | X_2)) P(X_1 > X_2)\}] \\
 &\quad - \mathbf{E}_{X_2} [\{\mathbf{E}_{X_1, Y_1} (\mathbf{E}_{Y_2} (Y_2 | X_1 < X_2) - \mathbf{E}_{Y_2} (Y_2 | X_2)) P(X_1 < X_2)\}] \quad (\text{第3形式})
 \end{aligned}$$

$d_{x,y,x}$ については若干複雑な形式となる。すなわち

$$\begin{aligned}
 (3.6) \quad d_{x,y,x} &= \mathbf{E}_{(X_1, Y_1), (X_2, Y_2)} \left[\{\text{sgn}(X_1 - X_2)\} \left| \begin{array}{cc} X_1 & X_2 \\ Y_1 & Y_2 \end{array} \right| \right] \quad (\text{原形式}) \\
 &= \mathbf{E}_{X_1, X_2} \left[\{\text{sgn}(X_1 - X_2)\} \left| \begin{array}{cc} X_1 & X_2 \\ \mathbf{E}(Y_1 | X_1) & \mathbf{E}(Y_2 | X_2) \end{array} \right| \right] \\
 &= \mathbf{E}_{X_2} \left\{ \left| \begin{array}{cc} 2\mathbf{E}(X_1 | X_1 > X_2) P(X_1 > X_2) - \mathbf{E}(X) & X_2 \\ 2\mathbf{E}(Y_1 | X_1 > X_2) P(X_1 > X_2) - \mathbf{E}(Y) & \mathbf{E}(Y_2 | X_2) \end{array} \right| \right\}
 \end{aligned}$$

従って結局

$$\begin{aligned}
 (3.7) \quad A_{x,y,x} &= 2E_{X_2} \left\{ \begin{array}{cc} E(X_1|X_1 > X_2) & X_2 \\ E(Y_1|X_1 > X_2) & E(Y_2|X_2) \end{array} \middle| P(X_1 > X_2) \right\} \\
 &\quad - E_{X_2} \left\{ \begin{array}{cc} E(X) & X_2 \\ E(Y) & E(Y_2|X_2) \end{array} \right\} \quad (\text{第1形式}) \\
 &= E_{X_2} \left\{ \begin{array}{cc} E(X) & X_2 \\ E(Y) & E(Y_2|X_2) \end{array} \right\} \\
 &\quad - 2E_{X_2} \left\{ \begin{array}{cc} E(X_1|X_1 < X_2) & X_2 \\ E(Y_1|X_1 < X_2) & E(Y_2|X_2) \end{array} \middle| P(X_1 < X_2) \right\} \quad (\text{第2形式}) \\
 &= E_{X_2} \left\{ \begin{array}{cc} E(X_1|X_1 > X_2) & X_2 \\ E(Y_1|X_1 > X_2) & E(Y_2|X_2) \end{array} \middle| P(X_1 > X_2) \right\} \\
 &\quad - E_{X_2} \left\{ \begin{array}{cc} E(X_1|X_1 < X_2) & X_2 \\ E(Y_1|X_1 < X_2) & E(Y_2|X_2) \end{array} \middle| P(X_1 < X_2) \right\} \quad (\text{第3形式})
 \end{aligned}$$

が得られる。

(2) 二変量に関する三次元 (体積) 平均差

二変量に関する三次元平均差 A_{xy} については (1) より多様な表現形式をとることが出来る。ここでは次節で直接用いる二表現を第1, 第2形式としてその定式を示すにとどめる。

$$\begin{aligned}
 (3.8) \quad A_{xy} &= E_{(X_1, Y_1)} E_{(X_2, Y_2)} E_{(X_3, Y_3)} \left[\text{sgn}(X_1 - X_3) \cdot \text{sgn}(Y_2 - Y_3) \right. \\
 &\quad \left. \times \begin{array}{ccc} 1 & 1 & 1 \\ X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \end{array} \right] \quad (\text{原形式}) \\
 &= E_{(X_3, Y_3)} \left(\begin{array}{ccc} 2P(X_1 > X_3) - 1 & & \\ 2P(X_1 > X_3)E(X_1|X_1 > X_3) - E(X), & & \\ 2P(X_1 > X_3)E(Y_1|X_1 > X_3) - E(Y), & & \\ & 2P(Y_2 > Y_3) - 1 & \\ & 2P(Y_2 > Y_3)E(X_2|Y_2 > Y_3) - E(X), & X_3 \\ & 2P(Y_2 > Y_3)E(Y_2|Y_2 > Y_3) - E(Y), & Y_3 \end{array} \right) \\
 &= 4 E_{(X_3, Y_3)} \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ E(X_1|X_1 > X_3) & E(X_2|Y_2 > Y_3) & X_3 \\ E(Y_1|X_1 > X_3) & E(Y_2|Y_2 > Y_3) & Y_3 \end{array} \middle| P(X_1 > X_3)P(Y_2 > Y_3) \right\} \\
 &\quad - 2 E_{(X_3, Y_3)} \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ E(X_1|X_1 > X_3) & E(X) & X_3 \\ E(Y_1|X_1 > X_3) & E(Y) & Y_3 \end{array} \middle| P(X_1 > X_3) \right\} \\
 &\quad - 2 E_{(X_3, Y_3)} \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ E(X) & E(X_2|Y_2 > Y_3) & X_3 \\ E(Y) & E(Y_2|Y_2 > Y_3) & Y_3 \end{array} \middle| P(Y_2 > Y_3) \right\}
 \end{aligned}$$

以上の計算を不等号の向きを逆転して計算すれば、二次元平均差の場合と同様に更に三通りの表現が可能である。然し差当り (3.8) 式を其等を代表する表現形式であるとして、次の表現

を第1形式とする.

$$(3.9) \quad A_{xy} = 4 \begin{matrix} E \\ (x, y) \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ E(X|X>x) & E(X|Y>y) & x & P(X>x)P(Y>y) \\ E(Y|X>x) & E(Y|Y>y) & y & \end{array} \right\} \\ - 2 \begin{matrix} E \\ x \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ E(X|X>x) & E(X) & x & P(X>x) \\ E(Y|X>x) & E(Y) & E(Y|x) & \end{array} \right\} \\ - 2 \begin{matrix} E \\ y \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ E(X) & E(X|Y>y) & E(X|y) & P(Y>y) \\ E(Y) & E(Y|Y>y) & y & \end{array} \right\}$$

ただし (X, Y) と (x, y) は同一分布に従うものとする. 又 $E(Y|x)$ 及び $E(X|y)$ はそれぞれ $X=x$ における Y の条件付期待値及び $Y=y$ における X の条件付期待値を表すものとする.

第1形式に対して計算手順を若干変更することにより, 次の第2形式を得ることが出来る. すなわち

$$(3.10) \quad A_{xy} = \begin{matrix} E \\ X_1 Y_2 \end{matrix} \begin{matrix} E \\ (X_3, Y_3) \end{matrix} \left[\left\{ \text{sgn}(X_1 - X_3) \text{sgn}(Y_2 - Y_3) \right\} \begin{array}{ccc|c} 1 & 1 & 1 & \\ X_1 & E(X_2|Y_2) & X_3 & \\ E(Y_1|X_1) & Y_2 & Y_3 & \end{array} \right] \\ = \begin{matrix} E \\ X_1 Y_2 \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & 1 - 2P(X_3 > X_1) - 2P(Y_3 > Y_2) & \\ X_1 & E(X_2|Y_2) & E(X_3) - 2E(X_3|X_3 > X_1)P(X_3 > X_1) & \\ E(Y_1|X_1) & Y_2 & E(Y_3) - 2E(Y_3|X_3 > X_1)P(X_3 > X_1) & \\ + 4P(X_3 > X_1, Y_3 > Y_2) & & & \\ - 2E(X_3|Y_3 > Y_2)P(Y_3 > Y_2) + 4E(X_3|X_3 > X_1, Y_3 > Y_2)P(X_3 > X_1, Y_3 > Y_2) & & & \\ - 2E(Y_3|Y_3 > Y_2)P(Y_3 > Y_2) + 4E(Y_3|X_3 > X_1, Y_3 > Y_2)P(X_3 > X_1, Y_3 > Y_2) & & & \end{array} \right) \\ = \begin{matrix} E \\ X_1 Y_2 \end{matrix} \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ X_1 & E(X_2|Y_2) & E(X_3) & \\ E(Y_1|X_1) & Y_2 & E(Y_3) & \end{array} \right) \\ - 2 \begin{matrix} E \\ X_1 Y_2 \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ X_1 & E(X_2|Y_2) & E(X_3|X_3 > X_1) & P(X_3 > X_1) \\ E(Y_1|X_1) & Y_2 & E(Y_3|X_3 > X_1) & \end{array} \right\} \\ - 2 \begin{matrix} E \\ X_1 Y_2 \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ X_1 & E(X_2|Y_2) & E(X_3|Y_3 > Y_2) & P(Y_3 > Y_2) \\ E(Y_1|X_1) & Y_2 & E(Y_3|Y_3 > Y_2) & \end{array} \right\} \\ + 4 \begin{matrix} E \\ X_1 Y_2 \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ X_1 & E(X_2|Y_2) & E(X_3|X_3 > X_1, Y_3 > Y_2) & P(X_3 > X_1, Y_3 > Y_2) \\ E(Y_1|X_1) & Y_2 & E(Y_3|X_3 > X_1, Y_3 > Y_2) & \end{array} \right\} \\ = \begin{array}{ccc|c} 1 & 1 & 1 & \\ E(X_1) & E(X_2) & E(X_3) & \\ E(Y_1) & E(Y_2) & E(Y_3) & \end{array} - 2 \begin{matrix} E \\ X_1 \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ X_1 & E(X_2) & E(X_3|X_3 > X_1) & P(X_3 > X_1) \\ E(Y_1|X_1) & E(Y_2) & E(Y_3|X_3 > X_1) & \end{array} \right\} \\ - 2 \begin{matrix} E \\ Y_2 \end{matrix} \left\{ \begin{array}{ccc|c} 1 & 1 & 1 & \\ E(X_1) & E(X_2|Y_2) & E(X_3|Y_3 > Y_2) & P(Y_3 > Y_2) \\ E(Y_1) & Y_2 & E(Y_3|Y_3 > Y_2) & \end{array} \right\}$$

$$+4E_{X_1} E_{Y_2} \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ X_1 & E(X_2|Y_2) & E(X_3|X_3>X_1, Y_3>Y_2) \\ E(Y_1|X_1) & Y_2 & E(Y_3|X_3>X_1, Y_3>Y_2) \end{array} \middle| P(X_3>X_1, Y_3>Y_2) \right\}$$

である、ここで X_1, Y_2 をそれぞれ x, y , X_i, Y_i をそれぞれ X, Y と書き変えると、結局第2形式として

$$(3.11) \quad A_{xy} = -2E_x \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ x & E(X) & E(X|X>x) \\ E(Y|x) & E(Y) & E(Y|X>x) \end{array} \middle| P(X>x) \right\} \\ -2E_y \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ E(X) & E(X|y) & E(X|Y>y) \\ E(Y) & y & E(Y|Y>y) \end{array} \middle| P(Y>y) \right\} \\ +4E_{x y} \left\{ \begin{array}{ccc} 1 & 1 & 1 \\ x & E(X|y) & E(X|X>x, Y>y) \\ E(Y|x) & y & E(Y|X>x, Y>y) \end{array} \middle| P(X>x, Y>y) \right\}$$

を得ることが出来る。

尚 $A_{y|x'}$ の計算は

$$(3.12) \quad Z = \frac{Y}{X}$$

として、まず X, Y に関する密度関数又は確率関数 $p(x, y)$ を X, Z のそれ $p(x, z)$ に変換し、更にこの分布より X に比例した加重密度又は確率関数 $p_w(x, z)$ を求め、その周辺分布 $p_w(z)$ について (3.2), (3.3) 又は (3.4) を適用するのが分かり易い。

4. パレート曲面の構造解析

二変量のパレート分布を解析する為に、次の三つの場合に大別することが出来る。すなわち

- (i) 二変量 X, Y が相互に独立である場合
- (ii) X の Y に関する及び Y の X に関する回帰が共に厳密に線形である場合
- (iii) 上述の回帰の一方又は双方が非線形となる一般の場合

である。もし周辺分布がそれぞれ

$$(4.1) \quad f(x; a, p) = pa^p/x^{p+1} ; \quad x > a > 0, p > 1 \\ = 0 ; \quad x \leq a, p > 1$$

及び

$$(4.2) \quad f(y; b, q) = qb^q/y^{q+1} ; \quad y > b > 0, q > 1 \\ = 0 ; \quad y \leq b, q > 1$$

である時、その周辺分布が (4.1) 及び (4.2) で与えられるという意味で、(i), (ii), (iii) に属するパレート分布としてはそれぞれ

$$(4.3) \quad f(x, y; a, b, p, q) = \frac{pqa^pb^q}{x^{p+1}y^{q+1}} ; \quad x > a > 0, y > b > 0; p > 1, q > 1 \\ = 0 ; \quad x \leq a, y \leq b; p > 1, q > 1$$

$$(4.4) \quad f(x, y; a, b, p) = \frac{p(p+1)ab^{p+1}}{(bx+ay-ab)^{p+2}} ; \quad x > a > 0, y > b > 0; p > 1 \\ x \leq a, y \leq b; p > 1 \quad (\text{Mardia 第1型})$$

(但し, この場合は $p=q>1$ に限られる)

及び

$$(4.5) \quad f(x, y; a, b, p, q, \alpha) = \frac{pq}{(1-\alpha)xy} \left\{ \left(\frac{a}{x} \right)^p \left(\frac{b}{y} \right)^q \right\}^{1/1-\alpha} \\ \cdot \left[\frac{2 \left\{ \alpha pq \log \frac{x}{a} \log \frac{y}{b} \right\}^{1/2}}{1-\alpha} \right]; \quad x > a > 0, y > b > 0; p > 1, q > 1, 0 < \alpha < 1 \\ = 0 \quad ; \quad x \leq a, y \leq b; p > 1, q > 1 \quad (\text{Mardia 第2型})$$

が知られている (Mardia [14] 参照).

以下, 逐次其等の解析を試みることにしよう. その際, 周辺分布 (4.1), (4.2) に対してはよく知られているように

$$(4.6) \quad E(X) = \frac{ap}{p-1}, \quad E(Y) = \frac{bq}{q-1}$$

$$(4.7) \quad \text{Var}(X) = \frac{a^2 p}{(p-1)^2(p-1)} \quad \text{但し } p > 2, \quad \text{Var}(Y) = \frac{b^2 q}{(q-1)^2(q-2)} \quad \text{但し } q > 2$$

$$(4.8) \quad A_x = \frac{2ap}{(p-1)(2p-1)}, \quad A_y = \frac{2bq}{(q-1)(2q-1)}$$

従って

$$(4.9) \quad G_x = \frac{1}{2p-1}, \quad G_y = \frac{1}{2q-1}$$

であることに留意する必要がある.

(1) 相互独立型パレート分布の解析

此の分布は, (4.5) に於て $\alpha=0$ とした場合にあたるが, 回帰の性質による分類の観点と後に行う複雑な計算をチェックする意味で, 独立に扱うことにした. (4.3) 式にみるように, 此の場合 X, Y は相互に独立であるから, 我々は (3.5), (3.6) 両式より直ちに

$$(4.10) \quad A_{y,x} = 0, \quad A_{x,y} = 0$$

$$(4.11) \quad A_{x,y,x} = E(Y) A_x, \quad A_{y,x,y} = E(X) A_y$$

を得ることが出来る. 従って (2.14) 式より

$$(4.12) \quad G_{x,y}^{(1)} = G_x, \quad G_{x,y}^{(2)} = 0$$

となる. 又 (2.10) 式及び Taguchi [18] II, Proposition 20 (p. 602) によって

$$(4.13) \quad G_{xy} = \frac{3}{4} G_x G_y.$$

故に又 (2.14), (2.18) 式により

$$(4.14) \quad A_{xy} = \frac{3}{4} A_x A_y$$

を得ることが出来る.

然し此等の性質は相互に独立な二変量 X, Y の分布に共通した性格であり, 此の限りではパレート分布としての特徴は見当たらない. さきに示した第 2.1 図は此の種のパレート分布の裾

を切断した分布のローレンツ曲面を示すものである。

(2) **Mardia 第1型パレート分布の解析**

ここでは (4.4) 式を対象とする。この時

$$(4.15) \quad E(Y|X=x) = b + \frac{bx}{ap}, \quad \rho_{xy} = \frac{1}{p}; \quad p > 2 \quad (\text{Mardia [14] 参照})$$

であるから $A_{y,x}$ について (3.5) の第1形式により、容易に次式が得られる。

$$(4.16) \quad \begin{aligned} A_{y,x} &= 2E_{X_2} \left[E_{X_1} \left(b + \frac{bX_1}{pa} \mid X_1 > X_2 \right) - \left(b + \frac{bX_2}{pa} \right) \left(\frac{a}{X_2} \right)^p \right] \\ &= \frac{2b}{pa} E_{X_2} \left[\{ E_{X_1} (X_1 \mid X_1 > X_2) - X_2 \} \left(\frac{a}{X_2} \right)^p \right] \\ &= \frac{b}{pa} A_x \end{aligned}$$

故に

$$(4.17) \quad \rho_{y,x} = \frac{A_{y,x}}{A_y} = \frac{1}{p}, \quad \text{同様に} \quad \rho_{x,y} = \frac{1}{p},$$

$$\text{故に} \quad \rho_{xy;\Delta} = \frac{1}{p}; \quad p > 1$$

つまりこの場合

$$(4.18) \quad \rho_{xy} = \rho_{xy;\Delta} = \frac{1}{p} \quad p > 2$$

が成立するが、 $\rho_{xy;\Delta}$ の存在領域は ρ_{xy} のそれより広いことになる。

$$(4.19) \quad \begin{aligned} A_{x,y,x} &= 2E \left\{ \left| \begin{array}{cc} \frac{p}{p-1} X & X \\ b + \frac{b}{a} \frac{X}{p-1} & b + \frac{bX}{ap} \end{array} \right| \left(\frac{a}{X} \right)^p \right\} - E \left\{ \left| \begin{array}{cc} \frac{pa}{p-1} & X \\ \frac{pb}{p-1} & b + \frac{bX}{ap} \end{array} \right| \right\} \\ &= 2b \frac{a^p}{p-1} E(X^{1-p}) + bE(X) - \frac{pab}{p-1} \\ &= 2 \frac{abp}{(p-1)(2p-1)} \\ &= bA_x \end{aligned}$$

となり (4.15) の回帰直線が (2.17) 式で示したように

$$(4.20) \quad E(Y|X=x) = \frac{A_{y,x}}{A_x} x + \frac{A_{x,y,x}}{A_x}$$

で表わされる事を示している。

さて、(4.4) 式の A_{xy} を計算する場合は (2.19) 式により先ず $f(x, y; 1, 1, p)$ について A_{xy} を算出し、その結果に ab を乗ずればよい。

今 $f(x, y; 1, 1, p)$ に対して (3.11) 式で示した第2形式を適用すると

$$(4.21) \quad A_{xy} = -2E_x \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ x & \frac{p}{p-1} & \frac{px}{p-1} & \frac{1}{x^p} \\ 1+\frac{x}{p} & \frac{p}{p-1} & 1+\frac{x}{p-1} & \end{array} \right) \\ -2E_y \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{p}{p-1} & 1+\frac{y}{p} & 1+\frac{y}{p-1} & \frac{1}{y^p} \\ \frac{p}{p-1} & y & \frac{py}{p-1} & \end{array} \right) \\ +4E_{xy} \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ x & 1+\frac{y}{p} & \frac{p}{p-1}(x+y-1)-(y-1) & \frac{1}{(x+y-1)^p} \\ 1+\frac{x}{p} & y & \frac{p}{p-1}(x+y-1)-(x-1) & \end{array} \right)$$

第1項及び第2項を展開すると容易に次式が得られる。

$$(4.22) \quad E_x \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ x & \frac{p}{p-1} & \frac{px}{p-1} & \frac{1}{x^p} \\ 1+\frac{x}{p} & \frac{p}{p-1} & 1+\frac{x}{p-1} & \end{array} \right) = E_y \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{p}{p-1} & 1+\frac{y}{p} & 1+\frac{y}{p-1} & \frac{1}{y^p} \\ \frac{p}{p-1} & y & \frac{py}{p-1} & \end{array} \right) \\ = 0$$

従って A_{xy} は第3項のみとなり結局

$$(4.23) \quad A_{xy} = 4E_{xy} \left\{ \frac{p+1}{p-1} \frac{1}{(x+y-1)^{p-1}} - \frac{(p+1)^2}{p^2} \frac{xy}{(x+y-1)^p} \right\} \\ = 4 \left\{ \frac{p+1}{p-1} \int_1^\infty \int_1^\infty \frac{p^2 dx dy}{(x+y-1)^{p-1} x^{p+1} y^{p+1}} - (p+1)^2 \int_1^\infty \int_1^\infty \frac{dx dy}{(x+y-1)^p x^p y^p} \right\}$$

である。従って又二次元集中度 G_{xy} は (2.18), (4.6) 両式により

$$(4.24) \quad G_{xy} = (p-1)^2 \left\{ \frac{p+1}{p-1} \int_1^\infty \int_1^\infty \frac{dx dy}{(x+y-1)^{p-1} x^{p+1} y^{p+1}} \right. \\ \left. - \frac{(p+1)^2}{p^2} \int_1^\infty \int_1^\infty \frac{dx dy}{(x+y-1)^p x^p y^p} \right\}$$

となる。此の結果を更に明確にする為に今

$$(4.25) \quad I_1 = \int_1^\infty \int_1^\infty \frac{1}{(x+y-1)^{p-1} x^{p+1} y^{p+1}} dx dy$$

として $z = \frac{1}{y-1}$ により変換すると結局

$$(4.26) \quad I_1 = \int_0^\infty \left\{ \int_1^\infty \frac{dx}{(1+zx)^{p-1} x^{p+1}} \right\} \frac{z^{2p-2}}{(1+z)^{p+1}} dz$$

を得ることが出来る。ここで付録の積分公式 (1) を適用すると結局

$$(4.27) \quad I_1 = \frac{1}{2p-1} \int_0^\infty \frac{Z^{p-1} {}_2F_1(p-1, 2p-1; 2p; -z^{-1})}{(1+z)^{p+1}} dz$$

となる。更に付録の Kummer の変換 (2) によると

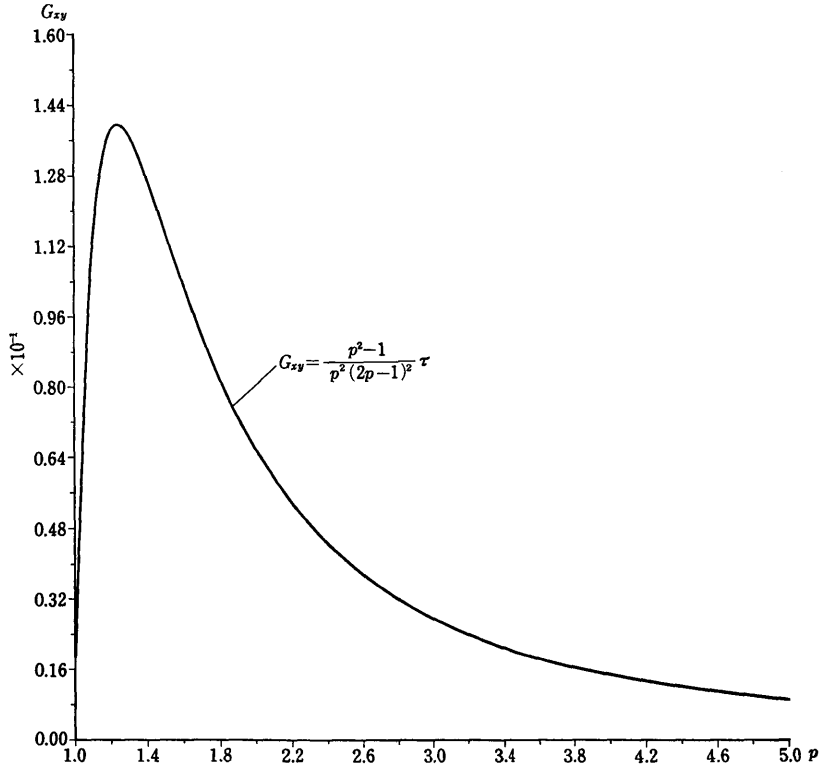
$$(4.28) \quad \begin{aligned} {}_2F_1(p-1, 2p-1; 2p; -z^{-1}) &= (1+z^{-1})^{-p+1} {}_2F_1\left(p-1, 1; 2p; \frac{z^{-1}}{1+z^{-1}}\right) \\ &= \frac{z^{p-1}}{(1+z)^{p-1}} {}_2F_1\left(p-1, 1; 2p; \frac{1}{1+z}\right) \end{aligned}$$

であるから

$$(4.29) \quad I_1 = \frac{1}{2p-1} \int_0^\infty \frac{z^{2p-2}}{(1+z)^{2p}} {}_2F_1\left(p-1, 1; 2p; \frac{1}{1+z}\right) dz$$

ここで $w = \frac{1}{1+z}$ とおくと結局

$$(4.30) \quad I_1 = \frac{1}{2p-1} \int_0^1 (1-w)^{2p-2} {}_2F_1(p-1, 1; 2p; w) dw$$



第 4.1 図 Pareto 係数 p に対する二次元集中度 G_{xy} の変化 (任意の a, b)

更に付録の積分公式 (3) を適用すると

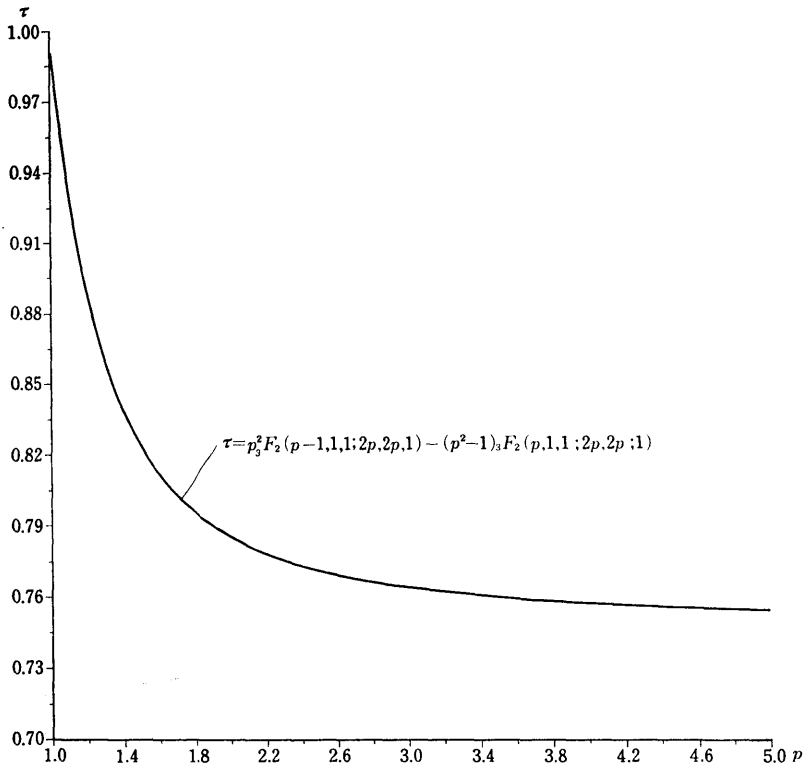
$$(4.31) \quad \begin{aligned} I_1 &= \frac{1}{2p-1} \frac{\Gamma(2p-1)}{\Gamma(2p)} {}_3F_2(p-1, 1, 1; 2p, 2p; 1) \\ &= \frac{1}{(2p-1)^2} {}_3F_2(p-1, 1, 1; 2p, 2p; 1) \end{aligned}$$

となる。他方

$$(4.32) \quad I_2 = \int_1^\infty \int_1^\infty \frac{1}{(x+y-1)^p x^p y^p} dx dy$$

は同様の手順を順次加えることにより I_1 に対応して

$$(4.33) \quad \begin{aligned} I_2 &= \int_0^\infty \left\{ \int_1^\infty \frac{dx}{(1+zx)^p x^p} \right\} \frac{z^{2p-2}}{(1+z)^p} dz \\ &= \frac{1}{2p-1} \int_0^\infty \frac{z^{p-2}}{(1+z)^{p-2}} {}_2F_1(p, 2p-1; 2p; -z^{-1}) dz \\ &= \frac{1}{2p-1} \int_0^\infty \frac{z^{2p-2}}{(1+z)^{2p}} {}_2F_1\left(p, 1; 2p; \frac{1}{1+z}\right) dz \end{aligned}$$



第4.2図 Pareto 係数 p に対する τ の変化

$$\begin{aligned}
 &= \frac{1}{2p-1} \int_0^1 (1-w)^{2p-2} {}_2F_1(p, 1; 2p; w) dw \\
 &= \frac{1}{(2p-1)^2} {}_3F_2(p, 1, 1; 2p, 2p, 1)
 \end{aligned}$$

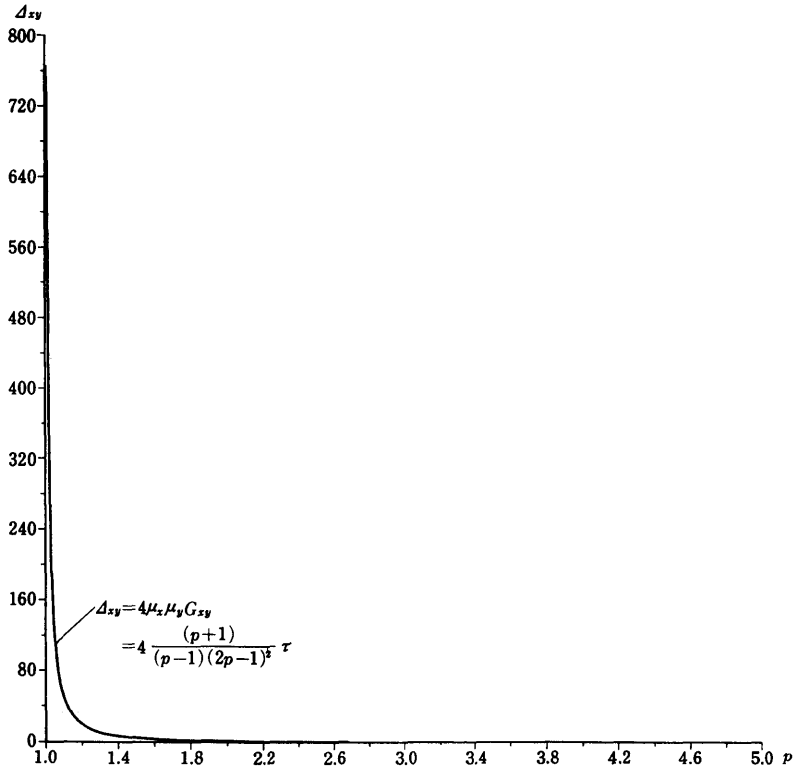
となる。従って(4.24)は結局

定理 4.1 「Mardia 第 1 型パレート分布 $f(x, y; a, b, p)$ (4.4) の二次元集中度 G_{xy} は

$$\begin{aligned}
 (4.34) \quad G_{xy} &= \frac{(p-1)^2}{(2p-1)^2} \left\{ \frac{p+1}{p-1} {}_3F_2(p-1, 1, 1; 2p, 2p; 1) \right. \\
 &\quad \left. - \frac{(p+1)^2}{p^2} {}_3F_2(p, 1, 1; 2p, 2p; 1) \right\} \\
 &= \frac{1 - \frac{1}{p^2}}{(2p-1)^2} \{ p^2 {}_3F_2(p-1, 1, 1; 2p, 2p; 1) \\
 &\quad - (p^2-1) {}_3F_2(p, 1, 1; 2p, 2p; 1) \}
 \end{aligned}$$

である。」

第 4.1 図~3 図は以上の計算に基づいて p に対応する集中度 G_{xy} , 及び平均差 A_{xy} の変化を図示したものである。



第 4.3 図 Pareto 係数 p に対する二次元平均差 A_{xy} の変化 ($a=b=1$)

$\frac{1}{p}$ は相関係数を与えるから、相関が高い程高い値を示す。グラフによると、平均差 A_{xy} の動きは、分散に対応するものとしてよく理解出来るが、集中度 G_{xy} は、平均値との関係で $p=1.2$ の近くで極値をもつことになり、興味のある課題を与えているといえよう。

又第4.2図は、 G_{xy} や A_{xy} の相関係数に対応する変化は、二次元分散のそれより緩慢であることを示している。

(3) Mardia 第2型パレート分布の解析

(4.5)式に関して Mardia はモメント相関 ρ_{xy} , 条件付密度関数 $f(x|y)$ 等につき次の諸式を与えている (Mardia [14] 参照*)。すなわち

$$(4.35) \quad \rho_{xy} = \alpha \{pq(p-2)(q-2)\}^{1/2} \{ (p-1)(q-1) - \alpha \}^{-1}; \quad p > 2, q > 2$$

$$(4.36) \quad f(x|y) = \frac{p}{(1-\alpha)x} \left\{ \left(\frac{a}{x} \right)^p \left(\frac{b}{y} \right)^{q\alpha} \right\}^{1/1-\alpha} \cdot I_0 \left[\frac{2 \left\{ \alpha pq \log \left(\frac{x}{a} \right) \log \left(\frac{y}{b} \right) \right\}^{1/2}}{1-\alpha} \right]$$

$$f(y|x) = \frac{q}{(1-\alpha)y} \left\{ \left(\frac{a}{x} \right)^{ap} \left(\frac{b}{y} \right)^q \right\}^{1/1-\alpha} \cdot I_0 \left[\frac{2 \left\{ \alpha pq \log \left(\frac{x}{a} \right) \log \left(\frac{y}{b} \right) \right\}^{1/2}}{1-\alpha} \right]$$

$$(4.37) \quad E(x|y) = \frac{pa}{p-1+\alpha} \left(\frac{y}{b} \right)^{q\alpha/(p-1+\alpha)}$$

$$E(y|x) = \frac{qb}{q-1+\alpha} \left(\frac{x}{a} \right)^{p\alpha/(q-1+\alpha)}$$

$$(4.38) \quad \text{Var}(x|y) = a^2 p \left[\frac{(b^{-1}y)^{2\alpha q/(2\alpha-2+p)}}{2\alpha-2+p} - \frac{p(b^{-1}y)^{2\alpha q/(\alpha-1+p)}}{(\alpha-1+p)^2} \right]; \quad p > 2$$

$$\text{Var}(y|x) = b^2 q \left[\frac{(a^{-1}x)^{2\alpha p/(2\alpha-2+q)}}{2\alpha-2+q} - \frac{q(a^{-1}x)^{2\alpha p/(\alpha-1+q)}}{(\alpha-1+q)^2} \right]; \quad q > 2$$

である。

(4.37)式によると x の y に関する回帰が線形ならば

$$(4.39) \quad q\alpha = p-1+\alpha \quad \text{すなわち} \quad \alpha = \frac{p-1}{q-1}$$

$0 \leq \alpha < 1$ だからこの時

$$(4.40) \quad \alpha \neq \frac{q-1}{p-1} \quad \text{すなわち} \quad p\alpha \neq q-1+\alpha$$

従って Mardia 第2型パレート分布の x の y , 又は y の x に関する回帰は、一方又は両方が非線形である。

さて、相互独立型 Mardia 第1型の場合と同様に、 $A_{y..}$ を計算する為に (3.5) 式の第1形式を適用すると

$$(4.41) \quad A_{y..} = 2E_{X_2} \left[\left\{ E_{X_1} \left(\frac{bq \left(\frac{X_1}{a} \right)^{p\alpha/(q-1+\alpha)}}{q-1+\alpha} \middle| X_1 > X_2 \right) - \frac{bq \left(\frac{X_2}{a} \right)^{p\alpha/(q-1+\alpha)}}{q-1+\alpha} \right\} P(X_1 > X_2) \right]$$

* $f(x|y)$, $E(x|y)$ 及び $\text{Var}(x|y)$ に関する Mardia の計算には、明らかな誤植が含まれている。

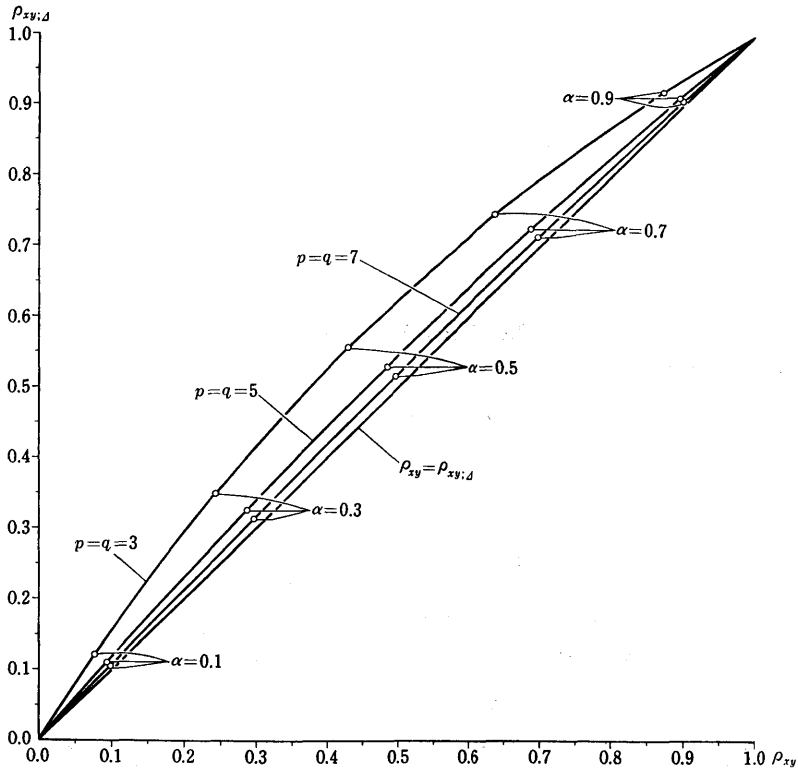
$$\begin{aligned}
 &= \frac{2bq}{q-1+\alpha} \mathbb{E}_{X_2} \left[\left\{ \frac{q-1+\alpha}{q-1} \left(\frac{X_2}{a} \right)^{p\alpha/(q-1+\alpha)} - \left(\frac{X_2}{a} \right)^{p\alpha/(q-1+\alpha)} \right\} \left(\frac{a}{X_2} \right)^p \right] \\
 &= \frac{2q}{q-1} \frac{b\alpha}{q-1+\alpha} \mathbb{E}_{X_2} \left\{ \left(\frac{X_2}{a} \right)^{p\alpha/(q-1+\alpha)-p} \right\} \\
 &= \frac{2\alpha q}{q-1} \cdot \frac{b}{2(q-1)+\alpha}
 \end{aligned}$$

従って

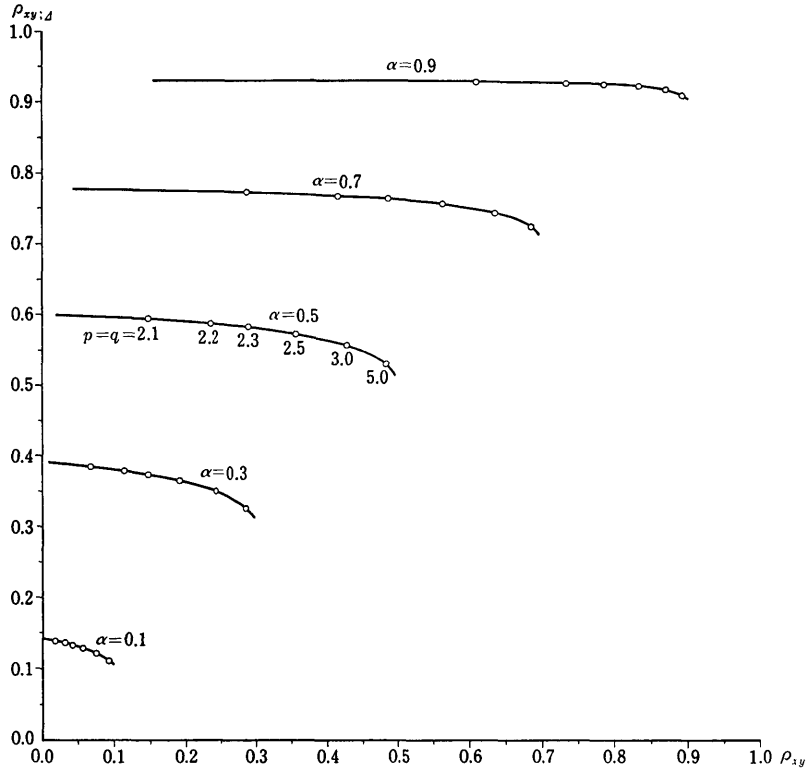
$$\begin{aligned}
 (4.42) \quad \rho_{y,x} &= \frac{A_{y,x}}{A_y} \\
 &= \alpha \frac{2q-1}{2(q-1)+\alpha}; \quad q > 1
 \end{aligned}$$

となる。故に又

$$\begin{aligned}
 (4.43) \quad \rho_{xy;\Delta} &= \sqrt{\rho_{y,x} \rho_{x,y}} \\
 &= \alpha \sqrt{\frac{2p-1}{2(p-1)+\alpha}} \sqrt{\frac{2q-1}{2(q-1)+\alpha}}; \quad p, q > 1
 \end{aligned}$$



第 4.4 図 $p=q$ を固定した場合、 α の変化に対応する ρ_{xy} と $\rho_{xy;\Delta}$ との関係
 $(\alpha = 0 \sim \frac{399}{400} \quad \alpha = 0.1, 0.3, 0.5, 0.7, 0.9 \text{ の各点に } \circ \text{ 記号})$



第 4.5 図 α を固定した場合, $p=q$ の変化に対応する ρ_{xy} と $\rho_{xy;A}$ との関係 ($p=q=2.1, 2.2, 2.3, 2.5, 3.0, 5.0$ の各点に○記号)

である。特に $\alpha=0$ のとき

$$(4.44) \quad \rho_{xy}, \rho_{xy;A}, \rho_{y,x}, \rho_{x,y} = 0$$

となるが, $\alpha \rightarrow 1$ のとき

$$(4.45) \quad \begin{aligned} \rho_{xy} &\neq 1 (p \neq q), \rho_{xy;A} = 1 \\ &= 1 (p = q) \end{aligned}$$

となり, 両者は必ずしも一致しないが, 一般に p, q が充分大ならば, 両者は殆ど一致する事は, 式の上からも明らかである。又 $\rho_{xy;A}$ は ρ_{xy} よりもその存在範囲が広い。両者が存在する範囲内で, α 又は $p=q$ が変化した場合の両者の関係を第 4.4 図, 第 4.5 図に示すことにしよう。

$\rho_{xy;A}$ は, p, q により影響されることが比較的少ないことが示されているといえよう。

次に (3.7) の第 1 形式に此の分布を適用すると, $a=b=1$ の時

$$(4.46) \quad A_{x,y,x} = 2E_x \left[\begin{array}{cc} \frac{px}{p-1} & x \\ \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)} & \frac{q}{q-1+\alpha} x^{p\alpha/(q-1+\alpha)} \end{array} \left(\frac{1}{x} \right)^p \right]$$

$$\begin{aligned}
 & -E_x \left[\begin{vmatrix} \frac{p}{p-1} & x \\ \frac{q}{q-1} & \frac{q}{q-1+\alpha} x^{p\alpha/(q-1+\alpha)} \end{vmatrix} \right] \\
 &= 2 \left(\frac{p}{p-1} \frac{q}{q-1+\alpha} - \frac{q}{q-1} \right) E_x (x^{p\alpha/(q-1+\alpha)+1-p}) \\
 & \quad + \frac{p}{p-1} \frac{q}{q-1} - \frac{p}{p-1} \frac{q}{q-1+\alpha} E_x (x^{p\alpha/(q-1+\alpha)}) \\
 &= 2 \left(\frac{p}{p-1} \frac{q}{q-1+\alpha} - \frac{q}{q-1} \right) \frac{p(q-1+\alpha)}{(2p-1)(q-1+\alpha) - p\alpha}
 \end{aligned}$$

となる。従って $\alpha=0$ すなわち X, Y が独立の場合

$$\begin{aligned}
 (4.47) \quad \Delta_{x,y,x} &= \frac{p}{p-1} \frac{q}{q-1} \frac{2}{2p-1} \\
 &= 2\mu_x \mu_y G_x = \mu_y \Delta_x
 \end{aligned}$$

となり、相互独立型パレート分布に関する (4.11) 式の結果と一致している。又 (4.46) 式より容易に

$$(4.48) \quad \alpha \cong \frac{q-1}{p-1}$$

に応じて

$$(4.49) \quad \Delta_{x,y,x} \cong 0$$

従って又 $\gamma_{y,x} \cong 0$ (第2節の比率係数) となることが理解出来る。

他方 (4.48), 従って (4.49) が成立するとき (4.37) の第二式の Y の X に関する回帰曲線のベキ指数 β は

$$(4.50) \quad \beta = \frac{p\alpha}{q-1+\alpha} \cong 1$$

となる。此の結果は又対数正規分布に関する (2.32), (2.33) 式の関係によく対応するものといえよう。

対数正規分布やパレート分布にみられる此等の関係は、より一般的な条件の下で成立することが証明出来る。(Taguchi [21] 参照)

二次元集中度 G_{xy} を計算する場合は Mardia 第1型の計算の場合と同様の理由により $f(x, y; 1, 1, p, q, \alpha)$ について計算すれば充分である。この計算に入る前に、一般化モメントに関する次の補助定理を証明しよう。

補助定理 4.1 「実数 s, t について

$$(4.51) \quad \frac{1}{1-\alpha} - \frac{s}{p} > 0$$

且つ

$$(4.52) \quad \left(\frac{1}{1-\alpha} - \frac{t}{q} \right) \left(\frac{1}{1-\alpha} - \frac{s}{p} \right) > \frac{\alpha}{(1-\alpha)^2}$$

ならば, Mardia 第2型パレート分布 $f(x, y; 1, 1, p, q, \alpha)$ (4.5), についての (s, t) 次モーメント $\mu_{st} = E_{(X, Y)}(X^s Y^t)$ は

$$(4.53) \quad \mu_{st} = \frac{1-\alpha}{\left\{1 - \frac{t}{q}(1-\alpha)\right\} \left\{1 - \frac{s}{p}(1-\alpha)\right\} - \alpha}$$

である。」

証明

$$(4.54) \quad \begin{aligned} \mu_{st} &= \int_a^\infty \int_b^\infty \frac{pq x^s y^t}{(1-\alpha)xy} \left[\left(\frac{a}{x} \right)^p \left(\frac{b}{y} \right)^q \right]^{1/(1-\alpha)} \\ &\quad \times I_0 \left[\frac{2 \left\{ \alpha pq \left(\log \frac{x}{a} \right) \left(\log \frac{y}{b} \right) \right\}^{1/2}}{1-\alpha} \right] dx dy \\ &\quad x > a > 0, \quad y > b > 0, \quad p, q > 1 \quad 0 \leq \alpha < 1 \end{aligned}$$

に於て

$$(4.55) \quad \begin{aligned} p \log \left(\frac{x}{a} \right) &= u^2 \\ q \log \left(\frac{y}{b} \right) &= v^2 \end{aligned}$$

とおけば, 結局

$$(4.56) \quad \begin{aligned} \mu_{st} &= \frac{4}{1-\alpha} a^s b^t \int_0^\infty \int_0^\infty uv \exp \left(- \left(\frac{1}{1-\alpha} - \frac{s}{p} \right) u^2 - \left(\frac{1}{1-\alpha} - \frac{t}{q} \right) v^2 \right) \\ &\quad \cdot I_0 \left\{ \frac{2\alpha^{1/2} uv}{1-\alpha} \right\} du dv \end{aligned}$$

となる. ここで (4.51) が成立すると付録の積分公式 (4) が適用可能となり結局

$$(4.57) \quad \begin{aligned} \mu_{st} &= \frac{4}{1-\alpha} a^s b^t \int_0^\infty v e^{-(1/(1-\alpha) - t/q)v^2} \\ &\quad \times \left\{ \frac{1}{2 \left(\frac{1}{1-\alpha} - \frac{s}{p} \right)} {}_1F_1 \left(1, 1; \frac{\alpha v^2}{(1-\alpha)^2 \left(\frac{1}{1-\alpha} - \frac{s}{p} \right)} \right) \right\} dv \\ &= \frac{2a^s b^t}{(1-\alpha) \left(\frac{1}{1-\alpha} - \frac{s}{p} \right)} \int_0^\infty v \exp \left[- \left\{ \frac{1}{1-\alpha} - \frac{t}{q} - \frac{\alpha}{(1-\alpha)^2 \left(\frac{1}{1-\alpha} - \frac{s}{p} \right)} \right\} v^2 \right] dv \end{aligned}$$

である. ここでもし (4.52) が成立すれば結局

$$(4.58) \quad \mu_{st} = \frac{a^s b^t}{(1-\alpha) \left\{ \left(\frac{1}{1-\alpha} - \frac{t}{q} \right) \left(\frac{1}{1-\alpha} - \frac{s}{p} \right) - \frac{\alpha}{(1-\alpha)^2} \right\}}$$

となる. ここで更に $a=b=1$ とすれば (4.53) 式が得られる (証明了).

さて, $f(x, y; 1, 1, p, q, \alpha)$ に対して (3.9) 式で示される第1形式を適用すると

$$(4.59) \quad A_{xy} = 4 \mathop{\text{E}}_{(x,y)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{px}{p-1} & \frac{p}{p-1} y^{q\alpha/(p-1+\alpha)} & x & (1/x^p)(1/y^q) \\ \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)} & \frac{qy}{q-1} & y & \end{array} \right)$$

$$- 2 \mathop{\text{E}}_x \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{px}{p-1} & \frac{p}{p-1} & x & \frac{1}{x^p} \\ \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)} & \frac{q}{q-1} & \frac{q}{q-1+\alpha} x^{p\alpha/(q-1+\alpha)} & \end{array} \right)$$

$$- 2 \mathop{\text{E}}_y \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{p}{p-1} & \frac{p}{p-1} y^{q\alpha/(p-1+\alpha)} & \frac{p}{p-1+\alpha} y^{q\alpha/(q-1+\alpha)} & \frac{1}{y^q} \\ \frac{q}{q-1} & \frac{q}{q-1} y & y & \end{array} \right)$$

第1項について

$$(4.60) \quad T_1 = \mathop{\text{E}}_{(x,y)} \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{px}{p-1} & \frac{p}{p-1} y^{q\alpha/(p-1+\alpha)} & x & (1/x^p)(1/y^q) \\ \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)} & \frac{q}{q-1} y & y & \end{array} \right)$$

$$= \mathop{\text{E}}_{(x,y)} \left[\left\{ \frac{-p}{p-1} \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)} y^{q\alpha/(p-1+\alpha)} \right. \right.$$

$$\left. \left. + \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)+1} + \frac{p}{p-1} y^{q\alpha/(p-1+\alpha)+1} \right. \right.$$

$$\left. \left. + \left(\frac{p}{p-1} \frac{q}{q-1} - \frac{p}{p-1} - \frac{q}{q-1} \right) xy \right\} \frac{1}{x^p y^q} \right]$$

$$= - \frac{p}{p-1} \frac{q}{q-1} \mu_{p\alpha/(q-1+\alpha)-p, q\alpha/(p-1+\alpha)-q} + \frac{q}{q-1} \mu_{q\alpha/(q-1+\alpha)-p+1, -q}$$

$$+ \frac{p}{p-1} \mu_{-p, q\alpha/(p-1+\alpha)-q+1} + \frac{-pq+p+q}{(p-1)(q-1)} \mu_{1-p, 1-q}$$

式中の一般化モーメントに補助定理 4.1 を適用すると

$$(4.61) \quad T_1 = - \frac{p}{p-1} \frac{q}{q-1} \frac{1-\alpha}{\left\{ 1 - \left(\frac{\alpha}{q-1+\alpha} - 1 \right) (1-\alpha) \right\}}$$

$$\frac{\left\{ 1 - \left(\frac{\alpha}{p-1+\alpha} - 1 \right) (1-\alpha) \right\}^{-\alpha}}{\left\{ 1 - \left(\frac{\alpha}{q-1+\alpha} - 1 + \frac{1}{p} \right) (1-\alpha) \right\} \{ 1 + (1-\alpha) \}^{-\alpha}}$$

$$\begin{aligned}
& + \frac{p}{p-1} \frac{1-\alpha}{\{1+(1-\alpha)\} \left\{ 1 - \left(\frac{\alpha}{p-1+\alpha} - 1 + \frac{1}{q} \right) (1-\alpha) \right\}} - \alpha \\
& + \frac{-pq+p+q}{(p-1)(q-1)} \frac{1-\alpha}{\left\{ 1 - \left(\frac{1}{p} - 1 \right) (1-\alpha) \right\} \left\{ 1 - \left(\frac{1}{q} - 1 \right) (1-\alpha) \right\}} - \alpha \\
= & - \frac{pq}{(p-1)(q-1)} \frac{(p-1+\alpha)(q-1+\alpha)}{\{2(p-1)+\alpha\} \{2(q-1)+\alpha\} - \alpha(p-1)(q-1)} \\
& + \frac{pq}{q-1} \frac{q-1+\alpha}{2p(q-1+\alpha) + (2-\alpha)\{(p-1)(q-1)-\alpha\}} \\
& + \frac{pq}{p-1} \frac{p-1+\alpha}{2q(p-1+\alpha) + (2-\alpha)\{(p-1)(q-1)-\alpha\}} \\
& + \frac{pq}{(p-1)(q-1)} \frac{-pq+p+q}{(2p-1)(2q-1) - \alpha(p-1)(q-1)}
\end{aligned}$$

となる。第2項に関しては

$$\begin{aligned}
(4.62) \quad T_2 &= \mathbf{E}_x \left[\begin{vmatrix} 1 & 1 & 1 \\ \frac{px}{p-1} & \frac{p}{p-1} & x \\ \frac{q}{q-1} x^{p\alpha/(q-1+\alpha)} & \frac{q}{q-1} & \frac{q}{q-1+\alpha} x^{p\alpha/(q-1+\alpha)} \end{vmatrix} \frac{1}{x^p} \right] \\
&= \mathbf{E}_x \left[\left[\left(\frac{q}{q-1} - \frac{p}{p-1} \frac{q}{q-1+\alpha} \right) x^{p\alpha/(q-1+\alpha)+1} \right. \right. \\
&\quad + \left(\frac{p}{p-1} \frac{q}{q-1+\alpha} - \frac{p}{p-1} \frac{q}{q-1} \right) x^{p\alpha/(q-1+\alpha)} \\
&\quad \left. \left. + \left(\frac{p}{p-1} \frac{q}{q-1} - \frac{q}{q-1} \right) x \right] \frac{1}{x^p} \right] \\
&= \left(\frac{q}{q-1} - \frac{p}{p-1} \frac{q}{q-1+\alpha} \right) \mu_{p\alpha/(q-1+\alpha)-p+1,0} \\
&\quad + \frac{pq}{p-1} \left(\frac{1}{q-1+\alpha} - \frac{1}{q-1} \right) \mu_{p\alpha/(q-1+\alpha)-p,0} \\
&\quad + \frac{q}{q-1} \left(\frac{p}{p-1} - 1 \right) \mu_{1-p,0}
\end{aligned}$$

であるが

$$\begin{aligned}
(4.63) \quad \mu_{s0} &= \int_1^\infty \frac{p}{x^{p-s+1}} dx \\
&= \frac{p}{p-s}
\end{aligned}$$

を考慮すると結局

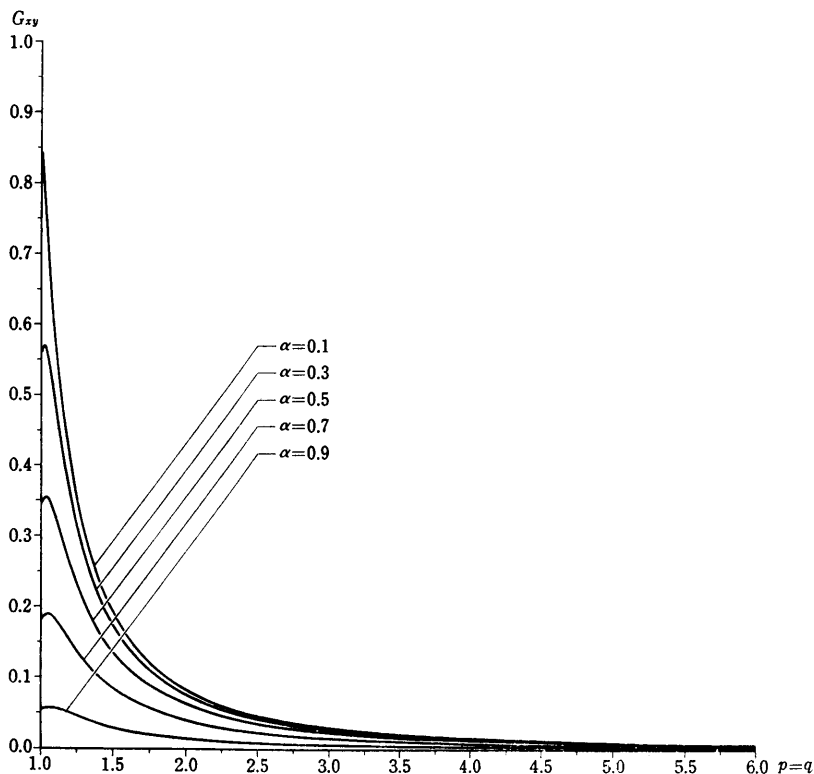
$$(4.64) \quad T_2 = \frac{pq}{(p-1)(q-1)} \frac{(p-1)\alpha - (q-1)}{(2p-1)(q-1) + (p-1)\alpha} - \frac{pq}{(p-1)(q-1)} \frac{\alpha}{(2q-1) + \alpha} + \frac{pq}{(p-1)(q-1)} \frac{1}{(2p-1)}$$

が得られる。同様にして第3項に関して

$$(4.65) \quad T_3 = E_y \left[\begin{array}{ccc|c} 1 & 1 & 1 & \\ \frac{p}{p-1} & \frac{p}{p-1} y^{q\alpha/(p-1+\alpha)} & \frac{p}{p-1+\alpha} y^{q\alpha/(p-1+\alpha)} & \frac{1}{y^q} \\ \frac{q}{q-1} & \frac{q}{q-1} y & y & \end{array} \right] \\ = \frac{pq}{(p-1)(q-1)} \cdot \frac{(q-1)\alpha - (p-1)}{(2q-1)(p-1) + (q-1)\alpha} - \frac{pq}{(p-1)(q-1)} \cdot \frac{\alpha}{2(p-1) + \alpha} + \frac{pq}{(p-1)(q-1)} \cdot \frac{1}{2p-1}$$

結局我々は次の定理を得ることが出来る。

定理 4.2 「Mardia 第2型パレート分布 $f(x, y; a, b, p, q, \alpha)$ (4.5) の二次元集中度 G_{xy} は

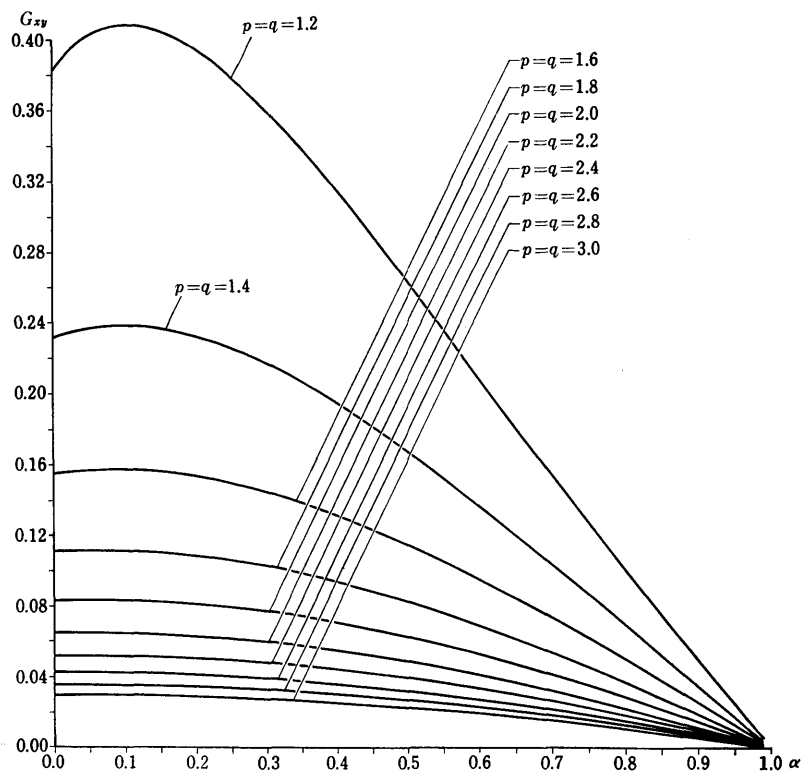


第 4.6 図 α を固定した場合、 $p=q$ の変化に対応する集中度 G_{xy}

$$\begin{aligned}
 (4.66) \quad G_{xy} &= \frac{4T_1 - 2T_2 - 2T_3}{4\mu_x\mu_y} \\
 &= - \frac{(p-1+\alpha)(q-1+\alpha)}{\{2(p-1)+\alpha\}\{2(q-1)+\alpha\} - \alpha(p-1)(q-1)} \\
 &\quad + \frac{(p-1)(q-1+\alpha)}{2p(q-1+\alpha) + (2-\alpha)\{(p-1)(q-1) - \alpha\}} \\
 &\quad + \frac{(q-1)(p-1+\alpha)}{2q(p-1+\alpha) + (2-\alpha)\{(p-1)(q-1) - \alpha\}} \\
 &\quad + \frac{-pq + p + q}{(2p-1)(2q-1) - \alpha(p-1)(q-1)} \\
 &\quad - \frac{1}{2} \left\{ \frac{(p-1)\alpha - (q-1)}{(2p-1)(q-1) + (p-1)\alpha} - \frac{\alpha}{2(q-1) + \alpha} + \frac{1}{2p-1} \right. \\
 &\quad \left. + \frac{(q-1)\alpha - (p-1)}{(2q-1)(p-1) + (q-1)\alpha} - \frac{\alpha}{2(p-1) + \alpha} + \frac{1}{2q-1} \right\}
 \end{aligned}$$

である。」

もし (4.39) が成立するならば $T_3=0$ となる。又もし $\alpha=0$ ならば容易に (4.66) から (4.13) 式



第4.7図 $p=q$ を固定した場合、 α の変化に対応する集中度 G_{xy}

を導くことが出来る。

系 1 「(4.5) 式に於て特に $p=q$ ならば Mardia 第 2 型対称パレート分布となるが、此の時は Y の X に関する及び X の Y に関する回帰は共に非線形である。且つこの場合

$$(4.67) \quad G_{xy} = - \frac{(p-1+\alpha)^2}{\{2(p-1)+\alpha\}^2 - \alpha(p-1)^2} \\ + \frac{2(p-1)(p-1+\alpha)}{2p(p-1+\alpha) + (2-\alpha)\{(p-1)^2 - \alpha\}} - \frac{p(p-2)}{(2p-1)^2 - \alpha(p-1)^2} \\ - \frac{\alpha-1}{2p-1+\alpha} + \frac{\alpha}{2(p-1)+\alpha} - \frac{1}{2p-1}$$

となる。」

第 4.6 図～第 4.7 図は (4.67) 式について α を固定した場合、 p による G_{xy} の変動及び p を固定した場合 α による G_{xy} の変動を示したものである。此等の何れの場合に於ても、Mardia 第 2 型パレート分布における G_{xy} の p による変動の場合と同様に単調な変化ではなく極値の存在が認められる。特に p, q を固定した場合 (4.7 図) $\alpha=0$ は XY の相互独立な条件を与えるが、平均値が一定であるにも拘わらず、集中度は此の点で最大とならない。従って G_{xy} は単なる散布度や残差とは区別される。それは不確実性又は多様性の測度とみるべきであろう。

5. 結 語

むすびとして、構造統計量の推定に関して一言触れなければならない。第 3 節で表現される諸構造母数 $A_x, A_y, A_{x,y}, A_{y,x}; A_{x,y,x}, A_{y,x,y}; A_{xy}$ 等は期待値記号 E を凡て標本平均の形式で置きかえることにより、標本構造統計量を得ることが出来る。其等の原形式は凡て不偏推定量であることは明らかである。又其等の分散については Lomnicki [10] の標本平均差の分散計算の方法が、他の統計量に関しても適用可能であると考えられる。然し乍ら其等に関する第 1, 第 2, 第 3 等の変形々式については Nair [16] の順序統計量としての標本平均差の表現とその分散計算が適切と思われる。又直接集中度及びその類似係数としての構造統計量 $G_x, G_y; G_{x,y}, G_{y,x}; G_{x,y,x}, G_{y,x,y}; G_{xy}$ を不偏推定しようとする場合は、Taguchi [20] の等確率及び逆回帰関数比例確率推定法の併用による二標本論推定の方式が適切と思われる。

以上は何れもノン・パラメトリックな推定法であるが、前二者の A_x の推定方式による誤差評価が何れも分散の存在を前提とするのに対して、第三の不偏推定方式はそれを必要としない点に於てユニークな方法であるといえる。

以上のように構造統計量の推定に関しては或程度の基本的足場が与えられているといえるが、此の概念自体が斬新であるから、その推定方式に関しては殆ど未知の段階であり、今後の研究成果に期待する他はない。

より重要な事は、構造統計量の推定に於ても、分散形式の精度評価が適当であるか否かという問題であり、更に遡ると、確率標本の妥当性が問われるのである。いわゆる統計推論に対して、パレート分布が投げかけた問題は依然として残されているように思われる。

謝 辞

積分公式の選択及びグラフのプログラムの作製の上での岡崎卓研究員の助言に感謝します。又原稿の整理、グラフ作製の全般について山田晴美補助員の助力であることを報告し感謝します。更に第 2.1 図の原形となったプラスチックモデルは、飯塚太美雄・道子御夫妻の苦心の勞

作であり、これをもとに 2.1 図を作図された小山田和枝嬢と共に改めて感謝を表明します。最後に計算を克明に点検されたレフェリー氏に深謝します。

付録：諸公式

$$(1) \int_b^{\infty} \frac{x^{s-1}}{(1+ax)^{\nu}} dx = a^{-\nu} b^{s-\nu} (\nu-s)^{-1} {}_2F_1(\nu, \nu-s; \nu-s+1; -a^{-1}b^{-1}); \operatorname{Re} s < \operatorname{Re} \nu$$

(Erdelyi 他 [4], p. 310, (21) 式)

$$(2) {}_2F_1(\alpha, \beta; \gamma; z) = (1-z)^{-\alpha} {}_2F_1\left(\alpha, \gamma-\beta; \gamma; \frac{z}{z-1}\right)$$

(Gradshteyn 他 [8], p. 1043, 森口他 [15], p. 59)

$$(3) \int_0^1 x^{\rho-1} (1-x)^{\sigma-1} {}_2F_1(\alpha, \beta; \gamma; x) dx$$

$$= \frac{\Gamma(\rho)\Gamma(\sigma)}{\Gamma(\rho+\sigma)} {}_3F_2(\alpha, \beta, \rho; \rho+\sigma; 1); \operatorname{Re} \rho > 0, \operatorname{Re} \sigma > 0, \operatorname{Re}(\gamma+\sigma-\alpha-\beta) > 0$$

(Gradshteyn 他 [8], p. 849, (5) 式)

$$(4) \int_0^{\infty} e^{-a^*x^2} x^{\nu-1} I_{\nu}(bx) dx$$

$$= \frac{\Gamma\left(\frac{\mu+\nu}{2}\right) b^{\nu}}{2^{\nu+1} a^{\mu+\nu} \Gamma(\nu+1)} {}_1F_1\left(\frac{\mu+\nu}{2}; \nu+1; \frac{b^2}{4a^2}\right); \operatorname{Re}(\mu+\nu) > 0$$

(森口他 [15], p. 200)

参考文献

- [1] 足利末男 (1966). 社会統計学史, 三一書房.
- [2] Arnold, B.C. (1981). *Pareto Distributions*, McDonald.
- [3] Blitz, R.C. and Brittain J.A. (1964). An extension of the Lorenz diagram to correlation of two variables, *Metron*, **23**, 137-143.
- [4] Erdelyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F.G. (1953). *Higher Transcendental Functions*, vol. I, II, McGraw-Hill Book Co., Inc., New York.
- [5] Engel, E. (1895). Die Lebenskosten belgischer Arbeiter — Familien früher und jetzt. (森戸辰男訳 (昭和 16 年). ベルギー労働者家族の生活費, 統計学古典選集 12, 大原社会問題研究所編, 栗田書店.)
- [6] Flaskämper, P. (1944). Allgemeine Statistik Grundriss der Statistick. (大橋隆憲, 足利末男訳 (昭和 28 年). 一般統計学, 統計学綱要 第一部, 農林統計協会刊.)
- [7] Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri, *Atti del R. Ist. veneto di Scienze, Lettere ed Arti*, **76**, parte II, Venezia.
- [8] Gradshteyn, I.S. and Ryzhik, I.M. (1965). *Table of Integrals, Series and Products* (Trans. ed. Alam Jeffry), Academic Press.
- [9] Johnson, N.L. and Kotz, S. (1970). *Continuous Univariate Distribution*, John-Wiley & Sons, New York.
- [10] Lomnicki, Z.A. (1952). The standard error of Gini's mean difference, *Ann. Math. Statist.*, **23**, 635-637.
- [11] Lorenz, M.O. (1965). *Publications of American Statistical Association*, Vol. 9.
- [12] Lunetta, G. (1972). Sulla concentrazione della distribuzioni doppie, *Società Italiana di Statistica*, **27**, Riunione Scientifica, Palermo, 127-150.
- [13] Lunetta, G. (1972). Di un indice di contrazione per variabili statistiche doppie, *Ann. Fac. Econ. Com. Catania*, A, **18**, 203-219.

- [14] Mardia, K.V. (1962). Multivariate Pareto distributions, *Ann. Math. Statist.*, **33**, 1008-1015.
- [15] 森口繁一, 宇田川銈久, 一松 信 (1957). 数学公式 III, 岩波全書.
- [16] Nair, U.S. (1936). The standard error of Gini's mean difference, *Biometrika*, **28**, 428-436.
- [17] Pareto, V.O. (1896-97). Cours d'economie politique, Vol III, Lausanne, p. 312.
- [18] Taguchi, T. (1972-73). I, II and III. On the two dimensional concentration surface and extension of concentration coefficient and Pareto distribution to the two dimensional case, *Ann. Inst. Statist. Math.*, **24**, 355-381, **25**, 599-619, **25**, 215-237.
- [19] 田口時夫 (1973). 二次元離散型分布の集中多面体, 集中係数及び新たな各種相関係数について, 統計数理研究所彙報, **20**, 77-115.
- [20] Taguchi, T. (1978). On an unbiased, consistent and asymptotically efficient estimation of Gini's concentration coefficient, *Metron*, **36**, 57-72.
- [21] Taguchi, T. (1981). On some properties of concentration surface and supersurface and its application to analysis of share distribution, *Research Memorandum* No. 206, The Institute of Statistical Mathematics.
- [22] Taguchi, T. (1981). On a multiple Gini's coefficient and some concentrative regressions, *Metron*, **39**, 69-98.
- [23] 田口時夫 (1982). 経済統計の解析概念について, 日本数学会統計数学分科会講演アブストラクト, p. 83.
- [24] 田口時夫 (1982). 統計的比率の多次元解析について, 科研費総合 A 「数理統計学の新しい展開」によるシンポジウム・多元データの解析予稿集.
- [25] 米沢治文 (1964). 地域分析のための統計指標の評価と整理, 日本統計学会々報, 42-47.

Concentration Analysis of Bivariate Paretoan Distribution
— A Proposition of Structural Statistics —

Tokio Taguchi

(The Institute of Statistical Mathematics)

In this paper, the bivariate Pareto distribution (Mardia [14]) are analyzed from a new point of view, that is multiple concentration analysis. The structural statistics is proposed by the author, corresponding to the geometrical properties of concentration surface (Taguchi [18]), which is an extension of Lorenz diagram. Traditional and extended Gini's coefficients, concentrative correlation and regression coefficients (Taguchi [22]) are all contained in this statistics. Finally, the author points out that the concentration coefficient is to be distinguished from the traditional dispersion measure and that it is to be a measure of uncertainty or multiformity.