



HOKKAIDO UNIVERSITY

Title	多項分布における母数空間の一つの拡張について
Author(s)	種市, 信裕; Taneichi, Nobuhiro; 佐藤, 義治 他
Citation	北海道大學工學部研究報告, 121, 59-67
Issue Date	1984-05-31
Doc URL	https://hdl.handle.net/2115/41877
Type	departmental bulletin paper
File Information	121_59-68.pdf



多項分布における母数空間の一つの拡張について

種 市 信 裕 佐 藤 義 治 河 口 至 商
(昭和58年12月27日受理)

On an Extension of the Parameter Space of the Multinomial Distribution

Nobuhiro TANEICHI, Yoshiharu SATO and Michiaki KAWAGUCHI
(Received December 27, 1983)

Abstract

The main purpose of this paper is to consider the geometrical properties of the parameter space with the metric based on α -entropy. For any value of α , it is shown that the parameter space of the multinomial distributions can be imbedded as a hypersurface in an Euclidean space. As a Riemannian hypersurface, the differential geometrical structures of the parameter space of the multinomial distributions are investigated.

With respect to the trinomial distributions, where the parameter space is regarded as a 2-dimensional hypersurface in 3-dimensional Euclidean space, several illustrations of the hypersurface depending on α are given and their features are considered.

1. は じ め に

Rao¹⁾によって、確率分布の母数空間に局所的に計量を導入することにより、確率分布の空間を幾何学(リーマン幾何学)的に解析することが提案され、この分野の研究が、吉沢²⁾、滝山³⁾、佐藤⁴⁾、又 Atkinson と Mitchell⁵⁾、そして Ingarden⁶⁾らによって行なわれている。又尾関⁷⁾は音声認識にリーマン構造を応用しようとした。

Čencov⁸⁾および Efron⁹⁾は確率分布の空間に接続の概念を導入しようとして試み、さらに甘利¹⁰⁾¹¹⁾¹²⁾¹³⁾¹⁴⁾や公文¹⁵⁾、長岡¹⁶⁾は、これらに α -接続と言う1パラメータ族を導入して、アファイン-接続空間まで理論を拡張し、統計学上有効な結果を多数導出している。

これらの研究においては、母数空間の基本計量テンソルとしては、Kullback¹⁷⁾の情報量の局所化としても得られる Fisher の情報行列が用いられている。

一方、確率分布間の相異度を表わす量として種々の divergence が考えられており、代表的なものとしては Kullback-Leibler¹⁷⁾の divergence がある。

Burbea と Rao¹⁸⁾は Shannon エントロピーの拡張である Havrda-Charvát の α -エントロピー¹⁹⁾による Jensen-difference として、divergence を定義している。(J-divergence)

さらに Rao²⁰⁾らは、この J-divergence を用いて、一般の確率分布の母数空間に計量を導入している。

この計量は、母数空間の計量であるから、母数の変換に対しては不変性を保つが、確率変数の変換に対しては不変性を保たない。しかしながらこの計量は α -エントロピー関数の性質により $\alpha=1$ の場合として、前出の研究において、母数空間の計量テンソルとして用いられているFisherの情報行列を含んでいることにより、より広範囲な幾何学的性質を調べることができる。

本研究は、多項分布の場合において、より広い意味での統計的推定論の基礎として、母数空間の幾何学的構造を次の2点によって明らかにした。

(i) n 項分布の母数空間におけるリーマンの曲率テンソルなどの種々の幾何学的量を求める。

さらにより具体的に

(ii) n 項分布の母数空間は $n-1$ 次元球面に、ある変数変換をほどこした n 次元超曲面として実現されることを示す。

加えて(ii)において、特に母数空間が2次元曲面として実現される3項分布について、 α に依存するこの曲面の α の変化による具体的形状と性質を考察した。

2. 確率分布の母数空間

可測空間 (X, \mathcal{B}) 上の確率測度が母数 $\theta = (\theta^1, \dots, \theta^n) \in \Theta$ (ただし Θ は \mathbf{R}^n の開部分集合とする。) に依在する場合、つまり任意の θ に対して定まる、ある σ -有限測度 $\nu(x)$ に絶対連続な確率測度 \mathbf{P}_θ の全体を考え、その確率密度関数の全体を $\mathbf{S} = \{\mathbf{p}(x, \theta)\}$ と表わす。

\mathbf{S} は同相写像 $\varphi: \{\mathbf{p}(x, \theta)\} \rightarrow \theta$ によって (\mathbf{S}, φ) を局所座標系とする多様体と見なせる。

この時、2つの確率分布 $\mathbf{p}(x, \theta_1)$, $\mathbf{p}(x, \theta_2)$ の間に擬距離 (divergence) を与え

$$(2.1) \quad \mathbf{D}(\theta_1, \theta_2) \equiv \mathbf{D}(\mathbf{p}(x, \theta_1), \mathbf{p}(x, \theta_2))$$

とすると $\mathbf{D}(\theta_1, \theta_2)$ は $\Theta \times \Theta \rightarrow \mathbf{R}$ への正定値符号の実数値関数であるが、一般に距離関数とはならず、距離関数の2乗のような性質を持つ。

しかし、 Θ の非常に近い2点 $\theta = (\theta^1, \dots, \theta^n)$, $\theta + d\theta = (\theta^1 + d\theta^1, \dots, \theta^n + d\theta^n)$ に対応する確率分布は、divergence $\mathbf{D}(\theta, \theta + d\theta)$ を適当に与えることによって Θ の各点の接空間に $d\theta^i$ の正定値二次形式として計算を導入することができる。

例えば、 $\mathbf{D}(\theta, \theta + d\theta)$ としてKullback-Leiblerのdivergence

$$(2.2) \quad \mathbf{D}(\theta, \theta + d\theta) = \int (\mathbf{p}(x, \theta) - \mathbf{p}(x, \theta + d\theta)) \log(\mathbf{p}(x, \theta) / \mathbf{p}(x, \theta + d\theta)) d\nu(x)$$

を与え、三次のオーダー以上の微小量を無視することによって

$$(2.3) \quad \mathbf{D}(\theta, \theta + d\theta) = g_{ij}(\theta) d\theta^i d\theta^j$$

ただし

$$\begin{aligned} g_{ij} &= \int \mathbf{p}(x, \theta) \left(\frac{1}{\mathbf{p}(x, \theta)} \frac{\partial \mathbf{p}(x, \theta)}{\partial \theta^i} \right) \left(\frac{1}{\mathbf{p}(x, \theta)} \frac{\partial \mathbf{p}(x, \theta)}{\partial \theta^j} \right) d\nu(x) \\ &= \mathbf{E} \left(\frac{\partial \mathbf{l}}{\partial \theta^i} \cdot \frac{\partial \mathbf{l}}{\partial \theta^j} \right) \quad (\mathbf{l} = \log \mathbf{p}(x, \theta)) \end{aligned}$$

と表わせる。すなわち $g_{ij}(\theta)$ は Fisher の情報行列である。

$g_{ij}(\theta)$ は、正定値で、二階の共変テンソルになっているので、確率密度関数の全体 $\mathbf{S} = \{\mathbf{p}(x, \theta)\}$ にリーマン計量を定義したことになり、母数空間をリーマン空間とみなすことができる。

一方Raoは、多項分布間の非類似度をHavrda-Charvátの α -エントロピーによるJensen

-difference の2倍として定義している。すなわち、 $\mathbf{p}=(p_1, \dots, p_n)$, $\mathbf{q}=(q_1, \dots, q_n)$ を2つの多項分布の母数とすると(2.4)式のように表わされる。

$$(2.4) \quad \mathbf{D}(\mathbf{p}, \mathbf{q}) = 2 \left\{ \mathbf{H}\left(\frac{\mathbf{p}+\mathbf{q}}{2}\right) - \frac{1}{2}\mathbf{H}(\mathbf{p}) - \frac{1}{2}\mathbf{H}(\mathbf{q}) \right\}$$

ただし

$$\mathbf{H}(\mathbf{p}) = \begin{cases} \frac{1}{\alpha-1} \left(1 - \sum_{i=1}^n p_i^\alpha \right) & (\alpha \neq 1) \\ - \sum_{i=1}^n p_i \log_e p_i & (\alpha = 1) \end{cases}$$

これは Burbea-Rao による多項分布間の J -divergence に対応している。この非類似度を divergence として一般の分布に拡張すると

$$(2.5) \quad \mathbf{D}(\theta_1, \theta_2) = \frac{1}{\alpha-1} \left\{ - \int \mathbf{p}\left(\mathbf{x}, \frac{\theta_1+\theta_2}{2}\right)^\alpha \mathbf{d}\nu(\mathbf{x}) + \frac{1}{2} \int \mathbf{p}(\mathbf{x}, \theta_1)^\alpha \mathbf{d}\nu(\mathbf{x}) + \frac{1}{2} \int \mathbf{p}(\mathbf{x}, \theta_2)^\alpha \mathbf{d}\nu(\mathbf{x}) \right\}$$

となる。したがってこの divergence において

$$(2.6) \quad \mathbf{D}(\theta, \theta+\mathbf{d}\theta) = \frac{1}{\alpha-1} \left\{ - \int \mathbf{p}\left(\mathbf{x}, \frac{2\theta+\mathbf{d}\theta}{2}\right)^\alpha \mathbf{d}\nu(\mathbf{x}) + \frac{1}{2} \int \mathbf{p}(\mathbf{x}, \theta)^\alpha \mathbf{d}\nu(\mathbf{x}) + \frac{1}{2} \int \mathbf{p}(\mathbf{x}, \theta+\mathbf{d}\theta)^\alpha \mathbf{d}\nu(\mathbf{x}) \right\}$$

より、三次のオーダー以上の微小量を見捨てることによって

$$(2.7) \quad \begin{aligned} \mathbf{D}(\theta, \theta+\mathbf{d}\theta) &= \frac{\alpha}{4} \left(\int \mathbf{p}(\mathbf{x}, \theta)^{\alpha-2} \frac{\partial \mathbf{p}(\mathbf{x}, \theta)}{\partial \theta^i} \frac{\partial \mathbf{p}(\mathbf{x}, \theta)}{\partial \theta^j} \mathbf{d}\nu(\mathbf{x}) \right) d\theta^i d\theta^j \\ &= \frac{\alpha}{4} \left(\int \mathbf{p}(\mathbf{x}, \theta)^\alpha \frac{\partial \mathbf{l}}{\partial \theta^i} \frac{\partial \mathbf{l}}{\partial \theta^j} \mathbf{d}\nu(\mathbf{x}) \right) d\theta^i d\theta^j \\ &\quad (\mathbf{l} = \log \mathbf{p}(\mathbf{x}, \theta)) \\ &= \frac{\alpha}{4} \mathring{g}_{ij} d\theta^i d\theta^j \end{aligned}$$

ここに

$$\mathring{g}_{ij} = \int \mathbf{p}(\mathbf{x}, \theta)^\alpha \frac{\partial \mathbf{l}}{\partial \theta^i} \frac{\partial \mathbf{l}}{\partial \theta^j} \mathbf{d}\nu(\mathbf{x})$$

\mathring{g}_{ij} は $\alpha=1$ の場合には Fisher の情報行列になるが、 $\alpha \neq 1$ の場合には基本的に異なっている。

またこの \mathring{g}_{ij} は θ の変換 $\bar{\theta}^\lambda = \bar{\theta}^\lambda(\theta^i)$ に対して

$$(2.8) \quad \mathring{g}_{\lambda\mu} = \frac{\partial \theta^i}{\partial \bar{\theta}^\lambda} \frac{\partial \theta^j}{\partial \bar{\theta}^\mu} \mathring{g}_{ij}$$

と言う変換を受け、二階の共変テンソルとなる。又 J -divergence $\mathbf{D}(\theta_1, \theta_2)$ は、 $(1 \leq \alpha \leq 2)$ の区間で convex になることが Burbea-Rao²⁰⁾により証明されているので $(1 \leq \alpha \leq 2)$ では、 $\mathbf{D}(\theta, \theta+\mathbf{d}\theta) = (\alpha/4) \mathring{g}_{ij} d\theta^i d\theta^j \geq 0$ となることから \mathring{g}_{ij} の正定値性が示される。故に $(1 \leq \alpha \leq 2)$ の場合には、 $(\alpha/4) \mathring{g}_{ij}(\theta)$ を用いて、各点の接空間にリーマン計量を導入できる。

3. 多項分布の母数空間

n 項分布

$$(3.1) \quad \mathbf{P}(x_1, \dots, x_n; \theta_1, \dots, \theta_{n-1}) = \frac{\left(\sum_{i=1}^n x_i\right)!}{x_1! x_2! \cdots x_n!} \theta_1^{x_1} \theta_2^{x_2} \cdots \theta_n^{x_n}$$

ただし

$$\theta_i \geq 0 (i=1, \dots, n-1), \theta_n = 1 - \sum_{i=1}^{n-1} \theta_i$$

の集合全体を $\mathbf{M} = \{\mathbf{p}(\mathbf{x}, \boldsymbol{\theta})\}$ と書くと (ただし, ここで n は固定されているものとする。) \mathbf{M} は同相写像 $\varphi: (\mathbf{x}, \boldsymbol{\theta}) \rightarrow (\theta_1, \dots, \theta_{n-1}) \in \Theta \subset \mathbf{R}^{n-1}$ によって (\mathbf{M}, φ) を局所座標系とする多様体と見なせる。そこで, 多項分布間の非類似度を前述した J -divergence $\mathbf{D}(\mathbf{p}, \mathbf{q})$ として, $\mathbf{D}(\mathbf{p}, \mathbf{q})$ によってこの多様体の任意の点における接平面に計量 (内積) を与えるために, 線素

$$(3.2) \quad ds^2 = \mathbf{D}(\boldsymbol{\theta}, \boldsymbol{\theta} + d\boldsymbol{\theta}) \\ = \frac{1}{\alpha-1} \left\{ \sum_{i=1}^n (\theta_i + d\theta_i)^\alpha + \sum_{i=1}^n \theta_i^\alpha - 2 \sum_{i=1}^n \left(\frac{(\theta_i + d\theta_i) + \theta_i}{2} \right)^\alpha \right\}$$

ただし

$$\sum_{i=1}^n \theta_i = 1, \theta_i \geq 0 (i=1, \dots, n), \sum_{i=1}^n d\theta_i = 0$$

を考える。ここで, 三次以上の微小量を無視すると

$$(3.3) \quad ds^2 = \mathbf{D}(\boldsymbol{\theta}, \boldsymbol{\theta} + d\boldsymbol{\theta}) \\ = \sum_{i=1}^{n-1} \frac{\alpha}{4} (\theta_i^{\alpha-2} + \theta_n^{\alpha-2}) (d\theta_i)^2 + \frac{\alpha}{2} \sum_{i \neq j}^{n-1} \theta_n^{\alpha-2} (d\theta_i)(d\theta_j) \\ = \sum_{i=1}^{n-1} \overset{\alpha}{g}_{ii} (d\theta_i)^2 + 2 \sum_{i \neq j}^{n-1} \overset{\alpha}{g}_{ij} (d\theta_i)(d\theta_j)$$

ただし

$$(3.4) \quad \begin{cases} \overset{\alpha}{g}_{ii} = \frac{\alpha}{4} (\theta_i^{\alpha-2} + \theta_n^{\alpha-2}) & (i=1, \dots, n-1) \\ \overset{\alpha}{g}_{ij} = \frac{\alpha}{4} (\theta_n) & (i \neq j, i, j=1, \dots, n-1) \end{cases}$$

一方, 行列 $(\overset{\alpha}{g}_{ij}) (i, j=1, \dots, n-1)$ の k 次 ($k=1, \dots, n-1$) 主座行列式は,

$$(3.5) \quad \left(\frac{\alpha}{4}\right)^k \left(\prod_{i=1}^k \theta_i\right)^{\alpha-2} \left\{ \sum_{i=1}^k \left(\frac{1}{\theta_i}\right)^{\alpha-2} + \left(\frac{1}{\theta_n}\right)^{\alpha-2} \right\}$$

で与えられる。したがって, $\alpha > 0$ ならば, 行列 $(\overset{\alpha}{g}_{ij})$ は正定値となるので, 多項分布の場合には, $\alpha > 0$ の場合に, 接平面における計算を, $\overset{\alpha}{g}_{ij}$ で与えることができる。故に, n 項分布の母数空間は $\overset{\alpha}{g}_{ij} (\alpha > 0)$ をリーマン計量とする $n-1$ 次元リーマン空間と見なすことができる。

ここで, 計量が (3.3) で与えられるリーマン空間について考えてみよう。一般に n 次元ユーク

クリッド空間の超曲面

$$(3.6) \quad \begin{cases} x_1 = \frac{1}{\sqrt{\alpha}} u_1^{\frac{\alpha}{2}} \\ x_2 = \frac{1}{\sqrt{\alpha}} u_2^{\frac{\alpha}{2}} \\ \vdots \\ x_{n-1} = \frac{1}{\sqrt{\alpha}} u_{n-1}^{\frac{\alpha}{2}} \\ x_n = \frac{1}{\sqrt{\alpha}} \left(1 - \sum_{i=1}^{n-1} u_i \right)^{\frac{\alpha}{2}} \end{cases}$$

を考えると、この空間の誘導計量テンソルは

$$(3.7) \quad \begin{cases} g_{ii} = \frac{\alpha}{4} (u_i^{\alpha-2} + u_n^{\alpha-2}) & (i=1, \dots, n-1) \\ g_{ij} = \frac{\alpha}{4} (u_n^{\alpha-2}) & (i \neq j, i, j=1, \dots, n-1) \end{cases}$$

ただし

$$u_n = 1 - \sum_{i=1}^{n-1} u_i$$

として与えられる。(3.7)と(3.4)を比較することによって、 $\theta_i (i=1, \dots, n-1)$ と $u_i (i=1, \dots, n-1)$ を同一視すると、 $\tilde{g}_{ij} = g_{ij} (i, j=1, \dots, n-1)$ となる。従ってこの対応により、この超曲面(3.6)と、 n 項分布の母数空間は等長となる。つまり、 n 項分布の母集団間の距離を $n-1$ 次元曲面

$$(3.8) \quad (\alpha x_1^2)^{\frac{1}{\alpha}} + (\alpha x_2^2)^{\frac{1}{\alpha}} + \dots + (\alpha x_n^2)^{\frac{1}{\alpha}} = 1 \\ (0 \leq x_i \leq 1/\sqrt{\alpha} \quad (i=1, \dots, n))$$

と等長な空間で測っていることになる。この曲面は、 $(\sqrt{\alpha} x_i)^{\frac{1}{\alpha}} = \xi_i (i=1, \dots, n)$ と変数交換することによって、 $\sum_{i=1}^n \xi_i^{\frac{1}{\alpha}} = 1 (0 \leq \xi_i \leq 1)$ と言う $n-1$ 次元球面の第一象限の部分となる。また(3.8)において、 $\alpha=1$ とすると、 $\sum_{i=1}^n x_i^2 = 1 (0 \leq x_i \leq 1)$ となり、Shannon エントロピーを用いた divergence による多項分布の母数空間は、 $n-1$ 次元球面の構造をもつ。一方 $\alpha=2$ とすると、 $\sum_{i=1}^n x_i = \sqrt{2}/2 (0 \leq x_i \leq \sqrt{2}/2)$ となり、多項分布の母数空間は、ユークリッド空間とみなすことができる。

次に \tilde{g}_{ij} を用いて、種々の幾何学的量を計算すると、以下ようになる。 (\tilde{g}_{ij}) の逆行列を (\tilde{g}^{ij}) とおくと

$$(3.9) \quad \begin{cases} \tilde{g}^{ij} = -\frac{4}{\alpha} \left\{ \frac{1}{(\theta_i \theta_j)^{\alpha-2} K} \right\} & (i \neq j, i, j=1, \dots, n-1) \\ \tilde{g}^{ii} = \frac{4}{\alpha} \sum_{\substack{j=1 \\ j \neq i}}^n \left\{ \frac{1}{(\theta_i \theta_j)^{\alpha-2} K} \right\} & (i=1, \dots, n-1) \end{cases}$$

ただし

$$K \equiv \sum_{i=1}^n (\theta_i^{2-\alpha})$$

クリストフェルの第一種記号

$$[ij, k] \equiv \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial \theta_j} + \frac{\partial g_{jk}}{\partial \theta_i} - \frac{\partial g_{ij}}{\partial \theta_k} \right)$$

を計算すると、次のようになる。

$$(3.10) \quad \begin{cases} [ii, i] = \frac{1}{8} \alpha (\alpha - 2) (\theta_i^{\alpha-3} - \theta_n^{\alpha-3}) \\ [ij, k] = -\frac{1}{8} \alpha (\alpha - 2) \theta_n^{\alpha-3} \quad (\text{その他}) \end{cases}$$

クリストフェルの第二種記号は、 $\left\{ \begin{smallmatrix} l \\ ij \end{smallmatrix} \right\} \equiv g^{lk} [ij, k]$ とおくと次のようになる。

$$(3.11) \quad \begin{cases} \left\{ \begin{smallmatrix} l \\ ij \end{smallmatrix} \right\} = -\frac{1}{2} (\alpha - 2) \left(\frac{1}{\theta_n \theta_i^{\alpha-2} K} \right) \quad (i \neq j) \\ \left\{ \begin{smallmatrix} l \\ ii \end{smallmatrix} \right\} = -\frac{1}{2} (\alpha - 2) \left(\frac{\theta_n + \theta_i}{\theta_n \theta_i \theta_i^{\alpha-2} K} \right) \\ \left\{ \begin{smallmatrix} i \\ ii \end{smallmatrix} \right\} = \frac{1}{2} (\alpha - 2) \left(\sum_{j=1}^n \frac{1}{\theta_i \theta_j^{\alpha-2} K} - \frac{1}{\theta_i^{\alpha-2} \theta_n K} \right) \end{cases}$$

ただし

$$K \equiv \sum_{i=1}^n (\theta_i^{2-\alpha})$$

リーマンの曲率テンソル

$$R_{hijk} \equiv \frac{\partial}{\partial \theta_i} [ik, h] - \frac{\partial}{\partial \theta_k} [ij, h] + \left\{ \begin{smallmatrix} l \\ ij \end{smallmatrix} \right\} [hk, l] - \left\{ \begin{smallmatrix} l \\ ik \end{smallmatrix} \right\} [hj, l]$$

の独立な成分は

$$(3.12) \quad \begin{cases} \textcircled{1} R_{hiji} = \frac{\alpha}{16} (\alpha - 2)^2 \left(\frac{1}{\theta_n \theta_i K} \right) \quad (h \neq j, i = k) \\ \textcircled{2} R_{hikh} = \frac{\alpha}{16} (\alpha - 2)^2 \left(\frac{1}{\theta_n \theta_n K} \right) \quad (h = j, i \neq k) \\ \textcircled{3} R_{hihj} = \frac{\alpha}{16} (\alpha - 2)^2 \left\{ \frac{(\theta_n + \theta_i + \theta_n)}{\theta_n \theta_n \theta_i K} \right\} \quad (h = j, i = k) \\ \textcircled{4} \text{その他} = 0 \end{cases}$$

ただし

$$K \equiv \sum_{i=1}^n (\theta_i^{2-\alpha})$$

これらの量は、 $\alpha = 1$ の場合、(3.9) は吉沢²⁾の結果の4倍(定数倍)、(3.10) と (3.12) は $\frac{1}{4}$ 倍(定数倍)、(3.11) は一致していることがわかる。

三項分布の場合には、 n 項分布の場合により具体的に次のようなことがわかる。三項分布間の計量テンソルは次のように与えられる。

$$(3.13) \quad \begin{bmatrix} \overset{\alpha}{g}_{11} & \overset{\alpha}{g}_{21} \\ \overset{\alpha}{g}_{12} & \overset{\alpha}{g}_{22} \end{bmatrix} = \frac{\alpha}{4} \begin{bmatrix} \theta_1^{\alpha-2} + \theta_3^{\alpha-2} & \theta_3^{\alpha-2} \\ \theta_3^{\alpha-2} & \theta_2^{\alpha-2} + \theta_3^{\alpha-2} \end{bmatrix}$$

ただし

$$\theta_3 = 1 - \theta_1 - \theta_2$$

この空間は 3 次元ユークリッド空間中の二次曲面

$$(3.14) \quad (\alpha x^2)^{\frac{1}{\alpha}} + (\alpha y^2)^{\frac{1}{\alpha}} + (\alpha z^2)^{\frac{1}{\alpha}} = 1 \quad (0 \leq x, y, z \leq 1/\sqrt{\alpha})$$

と同じ距離構造をもつ。この曲面の形状は α によりそれぞれ図 1 から図 5 のようになる。 $\alpha (> 0)$ が 1 より小さいと、外側に凸で立方体に近い形状をしている。 α が増加し、1 に近づくとつれてだんだん球面に近づき、 $\alpha = 1$ で (3.14) からわかるように、球面となる。また、 α がさらに増加して、 $\alpha = 2$ に近づくとつれてだんだん平坦になり、 $\alpha = 2$ で平面になる。 $\alpha = 2$ をこえると内側に凸になっていることがわかる。

このような性質は、この曲面の Gauss 曲率からもわかる。すなわち、 n 項分布の場合のリーマン曲率テンソルにより、三項分布の母数空間の Gauss 曲率を求めると

$$(3.15) \quad K \equiv R_{1212} / \det g \\ = \frac{(\theta - 2)^2 (\theta_1 \theta_2 \theta_3)^{\alpha - 3}}{\alpha \{ (\theta_1 \theta_2)^{\alpha - 2} + (\theta_2 \theta_3)^{\alpha - 2} + (\theta_3 \theta_1)^{\alpha - 2} \}^2}$$

ただし

$$\theta_3 = 1 - \theta_1 - \theta_2$$

となるので、 $\theta_1, \theta_2, \theta_3$ がすべて非負であることより、 $\alpha > 0$ の時には常に $K \geq 0$ であり、故に曲面は、各点の接平面に関して同一の側に存在することがわかる。

さらに、これらの曲面と、平面 $x=y, y=z$, および $z=x$ との交線は、それぞれこの曲面の測地線となっている。

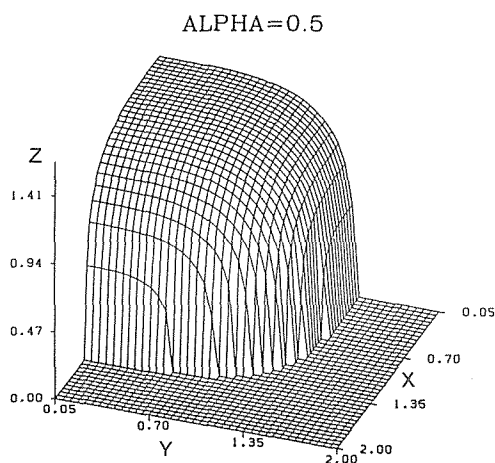


図 1 $\alpha = 0.5$ の場合

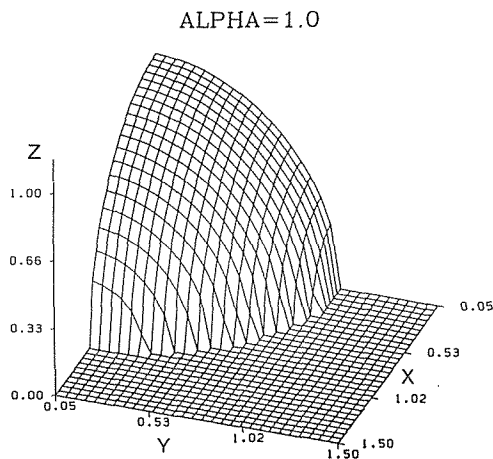
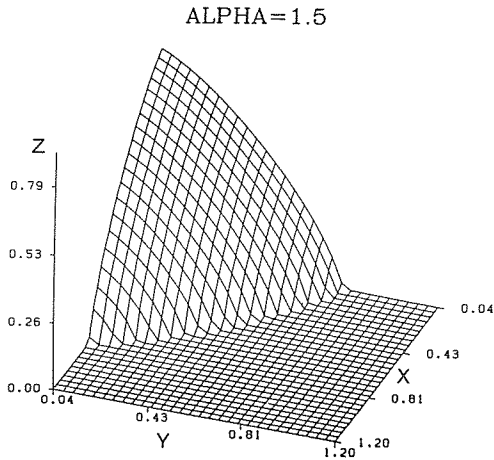
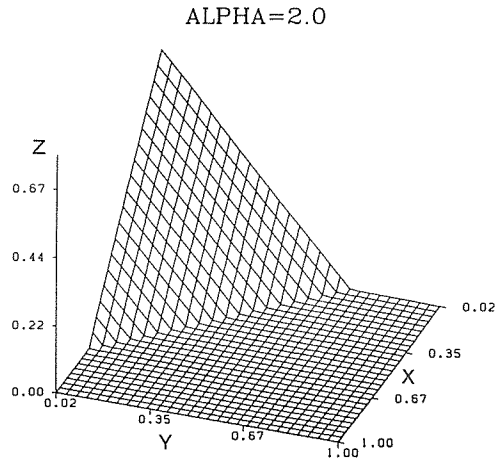
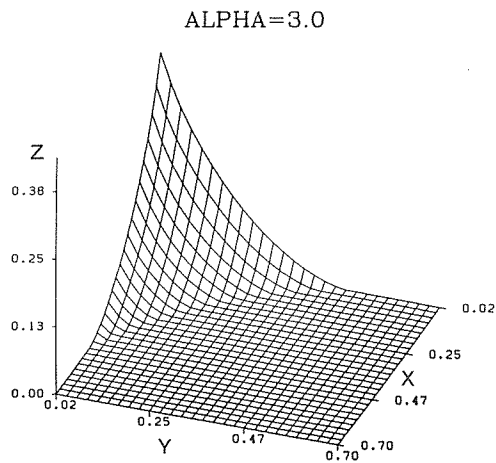


図 2 $\alpha = 1.0$ の場合

図3 $\alpha=1.5$ の場合図4 $\alpha=2.0$ の場合図5 $\alpha=3.0$ の場合

4. おわりに

n 項分布の J-divergence による母数空間は、 $n-1$ 次元曲面 $\sum_{i=1}^n (\alpha x_i^\alpha) = 1$ ($0 \leq x_i \leq 1/\sqrt{\alpha}$ ($i=1, \dots, n$)) と同じ距離構造を持つことがわかった, 特に, $\alpha=1$ の時は球面, $\alpha=2$ の時は平面と同じ距離構造を持つ定曲率空間である。他の場合は, 任意の点で異なる曲率を持っている。

$\alpha=1$ の場合の J-divergence をなすエントロピー関数は Shannon エントロピーであり, 又 $\alpha=2$ の場合のエントロピー関数は Gini-index²¹⁾²²⁾ と呼ばれ, 両方とも生態学などにおいて多様性の指数²³⁾として重要な関数である。これらによる J-divergence は, $\alpha=1$ の時に Lewontin²⁴⁾, $\alpha=2$ の時には Nei²⁵⁾によって, それぞれ遺伝学, 進化生態学において, 多項分布間の母集団間の差を計る量として“母集団間の多様性”と言う名で用いられている。つまり Lewontin は球面上

で差を測り、Nei は平面上で差を測っていることになる。これらのことから、実用上の面から言っても、 α による拡張はやみくもな拡張ではなく、これらを結ぶ重要な拡張となっているのである。

さらに Rao²⁶⁾²⁷⁾ は、分散分析 (ANOVA) の一般化の概念として、多様性の分析 (ANODIV) を提唱し、 α -エントロピーを用いた多様性の指数によって、カテゴリカルなデータにおける分割表の分析への可能性を述べており、今後、この分割表の分析などへの母数空間の幾何学的性質が応用できよう。

参考文献

- 1) Rao, C.R. : Bull. Calcutta Math. Soc., **37** (1945), pp. 81-91.
- 2) Yoshizawa, T. : Memorandum TYH-2, Harvard Univ. (1971), pp. 158-187.
- 3) 滝山竜三 : 電子通信学会論文誌, **57A** (1974) pp. 67-69.
- 4) Sato, Y., Sugawa, K. and Kawaguchi, M. : Rep. Math. Phys., **16** (1979), pp. 111-119.
- 5) Atkinson, C. and Mitchell, A.F.S. : Sankhyā A, **43** (1981), pp. 345-365.
- 6) Ingarden, R.S., Sato, Y., Sugawa, K. and Kawaguchi, M. : Tensor, n.s. **33** pp. 347-353.
- 7) 尾関和彦 : 音響学会研究会資料, **S 77-04** (1977) .
- 8) Čencov, N.N. : AMS, Rhode Island (1982).
- 9) Efron, B. : Ann. Statist., **3** (1975), pp. 1189-1242.
- 10) Amari, S. : POST RAAG Rep., 106 (1980).
- 11) Amari, S. : Ann. Statist., **10** (1982), pp. 357-385.
- 12) Amari, S. : Biometrika, **69** (1982), pp. 1-17.
- 13) Amari, S. and Kumon, M. : Ann. Statist. Math., **35** (1983), A, pp. 1-24.
- 14) 甘利俊一 : 電子通信学会論文誌 A, **66 A** (1983), pp. 492-499.
- 15) Kumon, M. and Amari, S. : Proc. Roy. Soc. (London), A-386 (1983).
- 16) Nagaoka, H. and Amari, S. : METR 82-7, Univ. Tokyo (1982).
- 17) Kullback, S.L. : Information Theory and Statistics, (1959), Wiley, New York.
- 18) Burbea, J. and Rao, C.R. : IEEE Trans. Information Theory, **28** (1982) pp. 489-495.
- 19) Havrda, J. and Charvát, F. : Kybernetika, **3** (1967), pp. 30-35.
- 20) Burbea, J. and Rao, C.R. : Jour. Multivariate Analysis, **12** (1982) pp. 575-596.
- 21) Gini, C. : Università di Cagliari III, Parte II (1912).
- 22) Simpson, E.H. : Nature, **163** (1949), pp. 688.
- 23) Pielou, E.C. : 数理生態学 (昭49), 産業図書.
- 24) Lewontin, R.C. : Evolutionary Biology, **6** (1972), pp. 381-398.
- 25) Nei, M. : Proc. Nat. Acad. Sci., **70** (1973) pp. 3321-3323.
- 26) Rao, C.R. : Tech. Rept. No. 80-10. Univ. of Pittsburgh (1980).
- 27) Rao, C.R. : Sankhyā A, **44** (1982), pp. 1-22.