数理解析研究所講究録 1452 巻 2005 年 63-77

エントロピー最大化原理と不等式

塚田　真 (Makoto TSUKADA)
Department of Information Sciences, Toho University
and
須鎗弘樹 (Hiroki SUYARI)
Department of Information and Image Sciences, Chiba University

1　はじめに

ここで取り扱うのは、定数 $s\neq 0,1$ に対する確率 $p=(p_1,\ldots,p_n)$ の関数 $F(p)=\sum_{i=1}^n {p_i}^s$ である。$s<0$ または $s>1$ のとき $F(p)$ は下に凸、$0<s<1$ のときは上に凸の関数となる。$F(p)$ は $p_1,\ldots,p_n$ の並べ替えによらないので、$p=(1/n,\ldots,1/n)$ で最小 ($s<0$ または $s>1$ のとき) 或いは最大 ($0<s<1$ のとき) となる。$s<0$ の場合は、上に非有界な関数で確率の全体の作る単体の境界で ∞ に発散する。特に関心があるのは、$e_1,\ldots,e_n,C\in\mathbb{R}$ が与えられたとき $\sum_{i=1}^n e_ip_i=C$ なる拘束条件の下での $F(p)$ の最小値問題 ($s<0$ または $s>1$ のとき) 或いは最大値問題 ($0<s<1$ のとき) である。拘束条件は n 次元空間の超平面 H をなすので、確率の全体の作る単体 Δ との公差する部分 $H\cap\Delta$ に p が属するときの $F(p)$ の最大最小問題である。拘束条件付き最大最小問題はラグランジェ未定乗数法を適用することができる。α,β を未定乗数として

$$L\left(p_1,\ldots,p_n,\alpha,\beta\right)=\sum_{i=1}^n {p_i}^s+\alpha\left(1-\sum_{i=1}^n p_i\right)+\beta\left(E-\sum_{i=1}^n e_ip_i\right)$$

を考える。

$$\frac{\partial L}{\partial p_i}=s{p_i}^{s-1}-\alpha-\beta e_i=0,\qquad i=1,\ldots,n$$

より

$$p_i=\left(\alpha+\beta e_i\right)^{1/(s-1)},\qquad i=1,\ldots,n$$

となる α および β がうまく見つかれば (α/s および β/s を改めて α および β で置き換えている)、これが解となる。しかしこのような α および β がいつでも存在するとは限らないし、存在が保証されたとしてもそれを見つけるのは厄介な作業である。

与えられた $q\neq 0$ および確率 $p_1,p_2,\ldots,p_n$ に対して、Rènyi エントロピーは

$$\frac{\log\sum_{i=1}^n {p_i}^q}{1-q}$$

で定義される。Tsallis エントロピーは

$$\frac{1-\sum_{i=1}^n {p_i}^q}{q-1}$$

で定義される。$-\log\sum_{i=1}^n {p_i}^q$ の $q=1$ における微分、および $-\sum_{i=1}^n {p_i}^q$ の $q=1$ における微分はいずれも $-\sum_{i=1}^n p_i\log p_i$、即ち Shannon エントロピーとなる。Rènyi エントロピー、Tsallis エン

トロピーはいずれもこの微分を差分に置き換えたものである。$\epsilon_1, \epsilon_2, \ldots, \epsilon_n \in \mathbb{R}$ および $C \in \mathbb{R}$ が与えられたとき、$\sum_{i=1}^{n} \epsilon_i p_i = C$ の条件の下で、Tsallis エントロピー或いは Rènyi エントロピーを最大にするという問題は、上で述べた問題に帰着される。

Tsallis エントロピー最大原理については、「$\frac{\sum_{i=1}^{n} \epsilon_i {p_i}^q}{\sum_{i=1}^{n} {p_i}^q} = C$ の条件の下で、Tsallis エントロピーを最大にせよ」という問題を考えることもある。この場合、

$$\mu_i = \frac{{p_i}^q}{\sum_{i=1}^{n} {p_i}^q} \qquad i = 1, 2, \ldots, n$$

とおく。そうすると、$\mu_1, \mu_2, \ldots, \mu_n \geq 0$ かつ $\mu_1 + \mu_2 + \cdots \mu_n = 1$ で $\sum_{i=1}^{n} \epsilon_i \mu_i = C$ であり、一方、$(\sum_{i=1}^{m} {p_i}^q)^{1/q} = \frac{1}{\sum_{i=1}^{m} \mu_i^{1/q}}$ の関係にあるので、目的関数を $\sum_{i=1}^{m} \mu_i^{1/q}$ として、$q > 1$ の場合は最大に、$q < 1$ の場合は最小にすればよい。$r = 1/q$ とすれば、問題は次のように書き換えられる。$\mu_1, \mu_2, \ldots, \mu_n \geq 0$ かつ $\mu_1 + \mu_2 + \cdots \mu_n = 1$ で $\sum_{i=1}^{n} \epsilon_i \mu_i = C$ の条件の下で $\sum_{i=1}^{m} \mu_i^r$ を、$r < 1$ の場合は最大に、$r > 1$ の場合は最小にせよ。結局、前節の問題を解けばよい。実際、この問題が

$$\mu_i = (\alpha + \beta e_i)^{1/(r-1)}, \qquad i = 1, 2, \ldots, n$$

という解を持てば、

$$\mu_i = (\alpha + \beta e_i)^{1/(1/q-1)} = (\alpha + \beta e_i)^{q/(1-q)}, \qquad i = 1, 2, \ldots, n$$

なので

$$p_i = {\mu_i}^{1/q} \left(\sum_{i=1}^{n} {p_i}^q \right)^{1/q} = \frac{(\alpha + \beta e_i)^{1/(1-q)}}{\sum_{i=1}^{n} (\alpha + \beta e_i)^{1/(1-q)}}, \qquad i = 1, 2, \ldots, n$$

が元の問題の解となる。α と β を置き直すことによって

$$p_i = (\alpha + \beta e_i)^{1/(1-q)}, \qquad i = 1, 2, \ldots, n$$

のような形が解であると言える。

連続版の Tsallis エントロピーは

$$\frac{1 - \int_{-\infty}^{+\infty} \phi(x)^q dx}{q - 1}$$

で定義される。Rènyi エントロピーは

$$\frac{\log \int_{-\infty}^{+\infty} \phi(x)^q dx}{1 - q}$$

で定義される。Shannon エントロピー最大原理の拘束条件をそのままに Tsallis エントロピーや Rènyi エントロピーを最大にする問題は興味ある問題である。q は定数であるから、目的関数は $\int_{-\infty}^{+\infty} \phi(x)^q dx$ を最大或いは最小にする問題に帰着できる。離散の場合と同様に、Tsallis エントロピー最大原理は拘束条件を

$$\frac{\int_{-\infty}^{+\infty} x^2 \phi(x) dx}{\int_{-\infty}^{+\infty} \phi(x)^q dx} = \sigma^2$$

で考えることも多い。この場合

$$\varphi(x) = \frac{\phi(x)^q}{\int_{-\infty}^{+\infty} \phi(x)^q dx}$$

とおくと

$$\int_{-\infty}^{+\infty} \varphi(x)^{1/q} dx = \frac{1}{\left(\int_{-\infty}^{+\infty} \phi(x)^q dx\right)^{1/q}}$$

の関係にあるので、結局 $\int_{-\infty}^{+\infty} \varphi(x)^{1/q} dx$ を最大或いは最小にする問題に帰着できる。即ち、

$$\begin{cases} \text{目的関数：} & \int_{-\infty}^{+\infty} \phi(x)^q dx \to \text{最大または最小} \\ \text{拘束条件：} & \phi \geq 0, \quad \int_{-\infty}^{+\infty} \phi(x) dx = 1, \quad \int_{-\infty}^{+\infty} x\phi(x) dx = 0, \quad \dfrac{\int_{-\infty}^{+\infty} x^2 \phi(x) dx}{\int_{-\infty}^{+\infty} \phi(x)^q dx} = \sigma^2 \end{cases}$$

という問題は、$r = 1/q$ として

$$\begin{cases} \text{目的関数：} & \int_{-\infty}^{+\infty} \varphi(x)^r dx \to \text{最小または最大} \\ \text{拘束条件：} & \varphi \geq 0, \quad \int_{-\infty}^{+\infty} \varphi(x) dx = 1, \quad \int_{-\infty}^{+\infty} x\varphi(x) dx = 0, \quad \int_{-\infty}^{+\infty} x^2 \varphi(x) dx = \sigma^2 \end{cases}$$

を解いて得られた最適解 φ に対して $\varphi^{1/q}$ を正規化した ψ が下の問題の最適解となる。実際、後者の問題の解が

$$\varphi(x) = \left(\alpha + \beta x^2\right)^{1/(r-1)} = \left(\alpha + \beta x^2\right)^{q/(1-q)}$$

であれば、前者の問題の解は

$$\varphi(x) = \left(\left(\alpha + \beta x^2\right)^{q/(1-q)}\right)^{1/q} = \left(\alpha + \beta x^2\right)^{1/(1-q)}$$

という形である (当然、α と β はそれぞれの問題で異なるが)。

2 不等式

以下に示す一連の不等式は $q = (q_1, \ldots, q_n)$ が与えられたとき、点 q を通るある超平面が存在して、その超平面の片側で q は $F(p)$ の最適解となっていることを述べたものである。

定理 1 (1) $s > 1$ とする。$p_1, \ldots, p_n, q_1, \ldots, q_n \geq 0$ に対して次の不等式が成立する。

$$\sum_{i=1}^{n} p_i {q_i}^{s-1} \geq \sum_{i=1}^{n} {q_i}^s \quad \Rightarrow \quad \sum_{i=1}^{n} {p_i}^s \geq \sum_{i=1}^{n} {q_i}^s$$

(2) $0 < s < 1$ とする。$p_1, \ldots, p_n \geq 0$ および , $q_1, \ldots, q_n > 0$ に対して次の不等式が成立する。

$$\sum_{i=1}^{n} p_i {q_i}^{s-1} \leq \sum_{i=1}^{n} {q_i}^s \quad \Rightarrow \quad \sum_{i=1}^{n} {p_i}^s \leq \sum_{i=1}^{n} {q_i}^s$$

(3) $s < 0$ とする。$p_1, \ldots, p_n \geq 0$ および , $q_1, \ldots, q_n > 0$ に対して次の不等式が成立する。

$$\sum_{i=1}^{n} p_i {q_i}^{s-1} \leq \sum_{i=1}^{n} {q_i}^s \quad \Rightarrow \quad \sum_{i=1}^{n} {p_i}^s \geq \sum_{i=1}^{n} {q_i}^s$$

系 1 $p_1, \ldots, p_n, q_1, \ldots, q_n \geq 0$ が

$$\sum_{i=1}^{n} p_i {q_i}^{s-1} = \sum_{i=1}^{n} {q_i}^s$$

を満たすとき、次の不等式が成立する

$$\sum_{i=1}^{n} {p_i}^s \geq \sum_{i=1}^{n} {q_i}^s, \qquad s>1 \text{ または } s<0 \text{ のとき}$$
$$\sum_{i=1}^{n} {p_i}^s \leq \sum_{i=1}^{n} {q_i}^s, \qquad 0<s<1 \text{ のとき}$$

補題 1 $D \subseteq \mathbb{R}$ とする。$f: D \to \mathbb{R}$ が微分可能な凸関数とするとき、$p_1, \ldots, p_n, q_1, \ldots, q_n \in D$ に対して次の不等式が成立する:

$$\sum_{i=1}^{n} \left(f(p_i) - f(q_i)\right) \geq \sum_{i=1}^{n} (p_i - q_i) f'(q_i).$$

証明 各 $i = 1, \ldots, n$ について $f(p_i) - f(q_i) \geq (p_i - q_i) f'(q_i)$ が成立することから明らか。■

この不等式を $f(x) = x \log x$ に対して適用すると、

$$\sum_{i=1}^{n} p_i \log p_i - \sum_{i=1}^{n} q_i \log q_i \geq \sum_{i=1}^{n} (p_i - q_i)(\log q_i + 1)$$

従って

$$\sum_{i=1}^{n} p_i \log p_i \geq \sum_{i=1}^{n} p_i \log q_i$$

を得る。これは Shannon の不等式に他ならない。この不等式の非可換版は、O. Klein の凸不等式と呼ばれる (D. Ruelle の *Statistical Mechancs* の pp.26-27)。

定理 1 の証明 補題を $s>1$ または $s<0$ のときは $f(x) = x^s$ とすれば

$$\sum_{i=1}^{n} \left({p_i}^s - {q_i}^s\right) \geq s \sum_{i=1}^{n} (p_i - q_i) {q_i}^{s-1}$$

が成立するので、$s>1$ のときは

$$\sum_{i=1}^{n} (p_i - q_i) {q_i}^{s-1} \geq 0 \quad \Rightarrow \quad \sum_{i=1}^{n} \left({p_i}^s - {q_i}^s\right) \geq 0$$

$s<0$ のときは

$$\sum_{i=1}^{n} (p_i - q_i) {q_i}^{s-1} \leq 0 \quad \Rightarrow \quad \sum_{i=1}^{n} \left({p_i}^s - {q_i}^s\right) \geq 0$$

を得る。また、$0<s<1$ の場合は $f(x) = -x^s$ とすれば

$$\sum_{i=1}^{n} \left({p_i}^s - {q_i}^s\right) \leq s \sum_{i=1}^{n} (p_i - q_i) {q_i}^{s-1}$$

が成立するので

$$\sum_{i=1}^{n} (p_i - q_i) {q_i}^{s-1} \leq 0 \quad \Rightarrow \quad \sum_{i=1}^{n} \left({p_i}^s - {q_i}^s\right) \leq 0$$

を得る。

3 離散の場合

$s \neq 0, 1$ および $e_1, \ldots, e_n, E \in \mathbb{R}$ に対して次の問題を考える。

$$(*) \begin{cases} \text{目的関数：} & \displaystyle\sum_{i=1}^{n} {p_i}^s \to \text{最大} \ (0 < s < 1) \ \text{または最小} \ (s < 0 \ \text{または} \ s > 1) \\ \text{拘束条件：} & p_1, \ldots, p_n \geq 0, \quad \displaystyle\sum_{i=1}^{n} p_i = 1, \quad \sum_{i=1}^{n} e_i p_i = E \end{cases}$$

$e_1 \leq e_2 \leq \cdots \leq e_n$ とする (このように仮定しても一般性を失わない)。実行可能解の集合がが空でないためには、$e_1 \leq E \leq e_n$ でなければならない。この条件を満たすとき、実行可能解の集合はコンパクトなので必ず最適解は存在する。特に、$E = e_1$ のときは $e_1 = \cdots = e_k < e_{k+1}$ を満たす k に対して

$$p_i = \begin{cases} 1/k, & i = 1, \ldots, k \\ 0, & i = k+1, \ldots, n \end{cases}$$

が最適解であり、また $E = e_n$ のときは $e_{n-k} < e_{n-k+1} = \cdots = e_n$ を満たす k に対して

$$p_i = \begin{cases} 0, & i = 1, \ldots n-k \\ 1/k, & i = n-k+1, \ldots, n \end{cases}$$

が最適解であることを見るのは易しい。従って、以下では $e_1 < E < e_n$ として話を進める。

定理 1 $\alpha, \beta \in \mathbb{R}$ が存在して、すべての $i = 1, \ldots, n$ に対して $\alpha + \beta e_i \geq 0$ であり、

$$q_i = (\alpha + \beta e_i)^{1/(s-1)}, \qquad i = 1, \ldots, n$$

である $q_1, \ldots, q_n$ が $(*)$ の実行可能解であるならば、それは最適解である。

証明 $p_1, \ldots, p_n$ を $(*)$ の任意の実行可能解とする。このとき、

$$\begin{aligned} \sum_{i=1}^{n} {q_i}^s &= \sum_{i=1}^{n} (\alpha + \beta e_i)^{s/(s-1)} = \sum_{i=1}^{n} (\alpha + \beta e_i)(\alpha + \beta e_i)^{1/(s-1)} \\ &= \sum_{i=1}^{n} (\alpha + \beta e_i) q_i = \alpha + \beta E = \sum_{i=1}^{n} (\alpha + \beta e_i) p_i = \sum_{i=1}^{n} {q_i}^{s-1} p_i \end{aligned}$$

が成立するので、第 3 節の系 1 より

$$\begin{aligned} \sum_{i=1}^{n} {p_i}^s &\geq \sum_{i=1}^{n} {q_i}^s, & s > 1 \ \text{または} \ s < 0 \ \text{のとき} \\ \sum_{i=1}^{n} {p_i}^s &\leq \sum_{i=1}^{n} {q_i}^s, & 0 < s < 1 \ \text{のとき} \end{aligned}$$

が導かれ、$q_1, \ldots, q_n$ が $(*)$ が最適解であることが証明された。■

定理 2 $s > 1$ とする。$\alpha, \beta \in \mathbb{R}$ および $\emptyset \neq I \subseteq \{1, \ldots, n\}$ が存在して、任意の $i \in I$ に対して $\alpha + \beta e_i \geq 0$ かつ任意の $i \notin I$ に対して $\alpha + \beta e_i < 0$ であり、

$$q_i = \begin{cases} (\alpha + \beta e_i)^{1/(s-1)}, & i \in I \\ 0, & i \notin I \end{cases}$$

である $q_1, \ldots, q_n$ が $(*)$ の実行可能解となるならば、それは最適解である。

証明 $p_1, \ldots, p_n$ を $(*)$ の任意の実行可能解とする。このとき、

$$\begin{aligned}\sum_{i=1}^{n} q_i{}^s &= \sum_{i\in I} q_i{}^s = \sum_{i\in I} (\alpha+\beta e_i)^{s/(s-1)} = \sum_{i\in I} (\alpha+\beta e_i)(\alpha+\beta e_i)^{1/(s-1)} \\ &= \sum_{i=1}^{n} (\alpha+\beta e_i)\, q_i = \alpha+\beta E = \sum_{i=1}^{n} (\alpha+\beta e_i)\, p_i \le \sum_{i=1}^{n} q_i{}^{s-1} p_i\end{aligned}$$

が成立するので、第3節の定理1により

$$\sum_{i=1}^{n} p_i{}^s \ge \sum_{i=1}^{n} q_i{}^s$$

導かれ、$q_1, \ldots, q_n$ が $(*)$ が最適解であることが証明された。■

上の一連の定理は、そのような実行可能解が存在するとすればそれは最適解であることを述べたものであり、存在については何も保証していない。以下ではその存在を証明するが、そのために準備が必要である。

$$f(x) = \frac{1}{Z}\sum_{i=1}^{n} e_i\,(x-e_i)^s, \qquad Z = \sum_{i=1}^{n} (x-e_i)^s$$

および

$$g(x) = \frac{1}{Z}\sum_{i=1}^{n} e_i\,(x+e_i)^s, \qquad Z = \sum_{i=1}^{n} (x+e_i)^s$$

を定義する。

定理3 $s>0$ のとき、f の定義域は $[e_n, \infty)$、g の定義域は $[-e_1, \infty)$ である。$s<0$ のとき、f の定義域は (e_n, ∞)、g の定義域は $(-e_1, \infty)$ であるが、

$$\lim_{x\to e_n} f(x) = e_n, \quad \lim_{x\to -e_1} g(x) = e_1$$

が成立する。

$s>0$ のとき f は単調増大で g は単調減少、$s<0$ のとき f は単調減少で g は単調増大減少である。

また、$s>0$ の場合 $s<0$ の場合のいずれも

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = \frac{e_1+\cdots+e_n}{n}$$

が成立する。

次の定義と補題は、Hardy-Littlewood-Pólya の *Inequalities*, p.43 による。

定義 $a_1, \ldots, a_n$ および $b_1, \ldots, b_n$ に対して

$$(a_i - a_j)(b_i - b_j) \ge 0 \quad \text{for all } i, j = 1, \ldots, n$$

が成立するとき、$a_1, \ldots, a_n$ と $b_1, \ldots, b_n$ は同順であるという。不等号の向きが逆のときは、逆順であるという。

もし、$a_1,\ldots,a_n$ と $b_1,\ldots,b_n$ の一方が整列しているならば、同順というのは他方も同順に整列されていて、逆順というのは他方が逆順に整列されていることに他ならない。

補題 1　$a_1,\ldots,a_n$ と $b_1,\ldots,b_n$ が同順であるとき、任意の $x_1,\ldots,x_n\geq 0$ に対して

$$\sum_{i=1}^n x_i\sum_{i=1}^n a_ib_ix_i\geq\sum_{i=1}^n a_ix_i\sum_{i=1}^n b_ix_i$$

が成立する。逆順の場合は、不等号の向きが逆となる。

証明

$$\begin{aligned}\sum_{i=1}^n x_i\sum_{i=1}^n a_ib_ix_i-\sum_{i=1}^n a_ix_i\sum_{i=1}^n b_ix_i&=\sum_{i,j=1}^n\left(x_ia_jb_jx_j-a_ix_ib_jx_j\right)\\ =\sum_{i,j=1}^n\left(x_ja_ib_ix_i-a_jx_jb_ix_i\right)&=\frac{1}{2}\sum_{i,j=1}^n\left(x_ia_jb_jx_j-a_ix_ib_jx_j+x_ja_ib_ix_i-a_jx_jb_ix_i\right)\\ =\frac{1}{2}\sum_{i,j=1}^n x_ix_j\left(a_jb_j-a_ib_j+a_ib_i-a_jb_i\right)&=\frac{1}{2}\sum_{i,j=1}^n x_ix_j\left(a_i-a_j\right)\left(b_i-b_j\right)\geq 0\end{aligned}$$

より示される。逆順の場合は最後の不等式の向きが逆となる。■

定理 3 の証明　定義域に関する最初の主張は、明らか。特に $s<0$ の場合の極限の式は、$f(x)$ 或いは $g(x)$ の分母分子を $x-e_n$ 或いは $x+e_1$ で割り算して、$x\to e_n$ 或いは $x\to -e_1$ とすれば得られる。

f および g の増減については、微分係数の正負を調べればよい。まず、f について見る。

$$Z'=s\sum_{i=1}^n\left(x-e_i\right)^{s-1}$$

より

$$f'(x)=\frac{s}{Z^2}\left(\sum_{i=1}^n e_i\left(x-e_i\right)^{s-1}\sum_{i=1}^n\left(x-e_i\right)^s-\sum_{i=1}^n e_i\left(x-e_i\right)^s\sum_{i=1}^n\left(x-e_i\right)^{s-1}\right)$$

である。ここで、$e_1,\ldots,e_n$ と $x-e_1,\ldots,x-e_n$ は逆順であることに注意する。補題より、

$$\sum_{i=1}^n e_i\left(x-e_i\right)^{s-1}\sum_{i=1}^n\left(x-e_i\right)^s\geq\sum_{i=1}^n e_i\left(x-e_i\right)^s\sum_{i=1}^n\left(x-e_i\right)^{s-1}$$

が言えるので、$s>0$ のときは f は単調増大、$s<0$ のときは f は単調減少であることが示される。g についても同様に、、$e_1,\ldots,e_n$ と $x+e_1,\ldots,x+e_n$ が同順であることから、補題に帰着される。最後の主張は、分母分子を x^s でわり算して $x\to\infty$ とすればよい。■

定理 4　$s>0$ のときは

$$\frac{\sum_{i=1}^n e_i\left(e_n-e_i\right)^s}{\sum_{i=1}^n\left(e_n-e_i\right)^s}\leq E\leq\frac{\sum_{i=1}^n e_i\left(e_i-e_1\right)^s}{\sum_{i=1}^n\left(e_i-e_1\right)^s}$$

$s<0$ のときは $e_1<E<e_n$ とする。このとき、$\alpha,\beta\in\mathbb{R}$ が存在して、任意の $i=1,\ldots,n$ に対して $\alpha+\beta e_i\geq 0$、かつ

$$\sum_{i=1}^n\left(\alpha+\beta e_i\right)^s=1.$$

証明

$$E=\frac{e_1+\cdots+e_n}{n}$$

のときは、$\alpha=n^{-1/s}$ とし $\beta=0$ とすればよい。$s>0$ の場合を考える。

$$\frac{\sum_{i=1}^n e_i\,(e_n-e_i)^s}{\sum_{i=1}^n (e_n-e_i)^s}\le E<\frac{e_1+\cdots+e_n}{n}$$

のとき、定理3により $g(x)=E$ となる x が存在する (中間値の定理)。そのとき

$$q_i=\frac{(x+e_i)^s}{Z}=(\alpha+\beta e_i)^s,\qquad i=1,\ldots,n$$

とすればよい。

$$\frac{e_1+\cdots+e_n}{n}<E\le\frac{\sum_{i=1}^n e_i\,(e_i-e_1)^s}{\sum_{i=1}^n (e_i-e_1)^s}$$

のときは、

$$f(x)=E$$

となる x が存在する (中間値の定理)。そのとき

$$q_i=\frac{(x-e_i)^s}{Z}=(\alpha+\beta e_i)^s,\qquad i=1,\ldots,n$$

とすればよい。

$s<0$ のときも同様である。■

系1　$s>1$ のときは

$$\frac{\sum_{i=1}^n e_i\,(e_n-e_i)^{1/(s-1)}}{\sum_{i=1}^n (e_n-e_i)^{1/(s-1)}}\le E\le\frac{\sum_{i=1}^n e_i\,(e_i-e_1)^s}{\sum_{i=1}^n (e_i-e_1)^s}$$

$s<1$ のときは $e_1<E<e_n$ とする。このとき、定理1を満たす $q_1,\ldots.q_n$ は必ず存在する。

定理5　$s>0$ とする。

$$e_1\le E<\frac{\sum_{i=1}^n e_i\,(e_n-e_i)^s}{\sum_{i=1}^n (e_n-e_i)^s}$$

または

$$\frac{\sum_{i=1}^n e_i\,(e_i-e_1)^s}{\sum_{i=1}^n (e_i-e_1)^s}<E\le e_n$$

であるとき、$\alpha,\beta\in\mathbb{R}$ および $\emptyset\ne I\subset\{1,\ldots,n\}$ が存在して、任意の $i\in I$ に対して $\alpha+\beta e_i\ge 0$、任意の $i\notin I$ に対して $\alpha+\beta e_i<0$、かつ

$$\sum_{i\in I}(\alpha+\beta e_i)^s=1.$$

証明

$$e_1\le E<\frac{\sum_{i=1}^n e_i\,(e_n-e_i)^s}{\sum_{i=1}^n (e_n-e_i)^s}$$

とする。$0\le k\le n-2$ に対して

$$f_k(x)=\frac{1}{Z_k}\sum_{i=1}^{n-k}e_i\,(x-e_i)^s$$

$$Z_k=\sum_{i=1}^{n-k}(x-e_i)^s$$

と定義する。このとき、f_k はすべて単調増加であり

$$f_0(e_n) = f_1(e_n) > f_1(e_{n-1}) = f_2(e_{n-1}) > f_2(e_{n-2}) = \\ \cdots = f_{n-2}(e_3) > f_{n-2}(e_2) = e_1$$

が言える。従って、

$$f_k(e_{n-k+1}) > E \geq f_k(e_{n-k})$$

を満たす $k \geq 1$ が唯一存在する。このとき、$E = f_k(x)$ となる $e_{n-k} \leq x < e_{n-k+1}$ が存在する。ここで、

$$q_i = \begin{cases} (x - e_i)^s / Z_k, & i = 1, \ldots, n-k \\ 0, & i = n-k+1, \ldots, n \end{cases}$$

とおけば

$$E = f_k(x) = \sum_{i=1}^{n} e_i q_i$$

であり

$$x - e_i < 0 \quad \text{for } i = n-k+1, \ldots, n$$

である。従って、$I = \{1, \ldots, n-k\}$ として

$$\alpha + \beta e_i = \frac{x - e_i}{{Z_k}^{1/s}}, \qquad i = 1, \ldots, n$$

とすればよい。

$$\frac{\sum_{i=1}^{n} e_i (e_i - e_1)^s}{\sum_{i=1}^{n} (e_i - e_1)^s} < E \leq e_n$$

であるときは $0 \leq k \leq n-2$ に対して

$$g_k(x) = \frac{1}{Z_k} \sum_{i=k+1}^{n} e_i (x + e_i)^s$$

$$Z_k = \sum_{i=k+1}^{n} (x + e_i)^s$$

と定義する。このとき、g_k はすべて単調減少であり

$$g_0(-e_1) = g_1(-e_1) < g_1(-e_2) = g_2(-e_2) < g_2(-e_3) = \\ \cdots = g_{n-2}(-e_{n-2}) < g_{n-2}(-e_{n-1}) = e_n$$

が言える。従って、

$$g_k(-e_k) < E \leq g_k(-e_{k+1})$$

を満たす $k \geq 1$ が唯一存在する。このとき、$E = g_k(x)$ となる $-e_k \geq x > -e_{k+1}$ が存在する。ここで、

$$q_i = \begin{cases} 0, & i = 1, \ldots, k \\ (x + e_i)^s / Z_k, & i = k+1, \ldots, n \end{cases}$$

とおけば

$$E = g_k(x) = \sum_{i=1}^{n} e_i q_i$$

であり

$$x + e_i < 0 \quad \text{for } i = 1, \ldots, k$$

である。従って、$I = \{k+1, \ldots, n\}$ として

$$\alpha + \beta e_i = \frac{x + e_i}{{Z_k}^{1/s}}, \qquad i = 1, \ldots, n$$

とすればよい。■

系 2　$s > 1$ とする。

$$e_1 \leq E < \frac{\sum_{i=1}^n e_i \left(e_n - e_i\right)^{1/(s-1)}}{\sum_{i=1}^n \left(e_n - e_i\right)^{1/(s-1)}}$$

または

$$\frac{\sum_{i=1}^n e_i \left(e_i - e_1\right)^{1/(s-1)}}{\sum_{i=1}^n \left(e_i - e_1\right)^{1/(s-1)}} < E \leq e_n$$

であるとき、定理 2 を満たす $q_1, \ldots. q_n$ は必ず存在する。

4　連続の場合

$$(*) \begin{cases} \text{目的関数:} & \displaystyle\int_{-\infty}^{+\infty} \phi(x)^s dx \to \text{最大}\ (0 < s < 1)\ \text{または最小}\ (s > 1) \\ \text{拘束条件:} & \phi \geq 0, \quad \displaystyle\int_{-\infty}^{+\infty} \phi(x) dx = 1, \quad \int_{-\infty}^{+\infty} x\phi(x) dx = 0, \quad \int_{-\infty}^{+\infty} x^2 \phi(x) dx = \sigma^2 \end{cases}$$

という問題を考える。この問題は $s < 0$ の場合に考えるのは無意味である。何故ならば、このとき目的関数は恒に発散してしまうからである。この問題の最適解を ψ とすると、$\psi(-\,\cdot\,)$ も最適解となる。目的関数は狭義凸或いは凹であるので、$\psi = \psi(-\,\cdot\,)$ でなければならない。即ち、最適解は原点に関して左右対称である。よって、次の問題を解けばよい。

$$(**) \begin{cases} \text{目的関数:} & \displaystyle\int_{0}^{+\infty} \phi(x)^s dx \to \text{最大}\ (0 < s < 1)\ \text{または最小}\ (s > 1) \\ \text{拘束条件:} & \phi \geq 0, \quad \displaystyle\int_{0}^{+\infty} \phi(x) dx = 1/2, \quad \int_{0}^{+\infty} x^2 \phi(x) dx = \sigma^2/2 \end{cases}$$

補題 1　$D \subseteq \mathbb{R}$ とする。$f : D \to \mathbb{R}$ が微分可能な凸関数で、$\phi, \psi : [0, +\infty) \to D$ であるとき、次の不等式が成立する:

$$\int_0^{+\infty} \left(f\left(\phi(x)\right) - f\left(\psi(x)\right)\right) dx \geq \int_0^{+\infty} \left(\phi(x) - \psi(x)\right) f'\left(\psi(x)\right) dx.$$

証明　$f\left(y_1\right) - f\left(y_2\right) \geq \left(y_1 - y_2\right) f'\left(y_2\right)$ が成立することから明らか。■

$s > 1$ のときは $f(x) = x^s$ とすれば補題より

$$\int_0^{+\infty} \left(\phi(x)^s - \psi(x)^s\right) dx \geq s \int_0^{+\infty} \left(\phi(x) - \psi(x)\right) \psi(x)^{s-1} dx$$

が成立するので、

$$\int_0^{+\infty}\left(\phi(x)-\psi(x)\right)\psi(x)^{s-1}dx \geq 0 \quad \Rightarrow \quad \int_0^{+\infty}\left(\phi(x)^s-\psi(x)^s\right)dx \geq 0$$

を得る。また、$0<s<1$ の場合は $f(x)=-x^s$ とすれば

$$\int_0^{+\infty}\left(\phi(x)^s-\psi(x)^s\right)dx \leq s\int_0^{+\infty}\left(\phi(x)-\psi(x)\right)\psi(x)^{s-1}dx$$

が成立するので

$$\int_0^{+\infty}\left(\phi(x)-\psi(x)\right)\psi(x)^{s-1}dx \leq 0 \quad \Rightarrow \quad \int_0^{+\infty}\left(\phi(x)^s-\psi(x)^s\right)dx \leq 0$$

を得る。定理としてまとめておこう。

定理 1

(1)　$s>1$ のとき

$$\int_0^{+\infty}\phi(x)\psi(x)^{s-1}dx \geq \int_0^{+\infty}\psi(x)^s dx \quad \Rightarrow \quad \int_0^{+\infty}\phi(x)^s dx \geq \int_0^{+\infty}\psi(x)^s dx$$

(2)　$0<s<1$ のとき

$$\int_0^{+\infty}\phi(x)\psi(x)^{s-1}dx \leq \int_0^{+\infty}\psi(x)^s dx \quad \Rightarrow \quad \int_0^{+\infty}\phi(x)^s dx \leq \int_0^{+\infty}\psi(x)^s dx$$

系 1

$$\int_0^{+\infty}\phi(x)\psi(x)^{s-1}dx = \int_0^{+\infty}\psi(x)^s dx$$

であるならば、

$$\int_0^{+\infty}\phi(x)^s dx \geq \int_0^{+\infty}\psi(x)^s dx, \qquad s>1 \text{ のとき}$$
$$\int_0^{+\infty}\phi(x)^s dx \leq \int_0^{+\infty}\psi(x)^s dx, \qquad 0<s<1 \text{ のとき}$$

定理 2　$\alpha,\beta\in\mathbb{R}$ が存在して、すべての $x\geq 0$ に対して $\alpha+\beta x\geq 0$ であり、

$$\psi(x)=\left(\alpha+\beta x^2\right)^{1/(s-1)}, \qquad x\in[0,\infty)$$

である ψ が $(**)$ の実行可能解であるならば、それは最適解である。

証明　ϕ を $(**)$ の任意の実行可能解とする。このとき、

$$\int_0^{+\infty}\psi(x)^s dx = \int_0^{+\infty}\left(\alpha+\beta x^2\right)^{s/(s-1)}dx = \int_0^{+\infty}\left(\alpha+\beta x^2\right)\left(\alpha+\beta x\right)^{1/(s-1)}dx$$
$$= \int_0^{+\infty}\left(\alpha+\beta x^2\right)\psi(x)dx = \frac{\alpha}{2}+\frac{\beta\sigma^2}{2} = \int_0^{+\infty}\left(\alpha+\beta x^2\right)\phi(x)dx = \int_0^{+\infty}\psi(x)^{s-1}\phi(x)dx$$

が成立するので、系 1 より

$$\int_0^{+\infty}\phi(x)^s dx \geq \int_0^{+\infty}\psi(x)^s dx, \qquad s>1 \text{ または } s<0 \text{ のとき}$$
$$\int_0^{+\infty}\phi(x)^s dx \leq \int_0^{+\infty}\psi(x)^s dx, \qquad 0<s<1 \text{ のとき}$$

が導かれ、ψ が (**) が最適解であることが証明された。■

この定理に出てくる ψ の形をした関数は、$s > 1$ のとき決して確率密度関数になりえない。また、$0 < s < 1$ の場合は $\alpha, \beta > 0$ でなければならないこともすぐわかる。そこで、ψ を

$$\psi(x) = \frac{1}{Z}(\gamma + x^2)^{1/(s-1)}, \qquad x \geq 0$$
$$\gamma > 0, \quad Z = 2\int_0^{+\infty} (\gamma + x^2)^{1/(s-1)} dx$$

のように書き換えることができる。ここで Z は必ず収束する。実際、ベータ関数の基本公式 (岩波全書「数学公式 1」p.222)

$$\int_0^{+\infty} \frac{dx}{x^\alpha (1+x^\lambda)^\beta} = \frac{1}{\lambda} B\left(\beta - \frac{1-\alpha}{\lambda}, \frac{1-\alpha}{\lambda}\right), \quad (\alpha < 1, \lambda, \beta > 0, \lambda\beta > 1-\alpha)$$

より

$$\int_0^{+\infty} \frac{dx}{(1+x^2)^{1/(1-s)}} = \frac{1}{2} B\left(\frac{1}{1-s} - \frac{1}{2}, \frac{1}{2}\right)$$

を得、更に

$$\int_0^{+\infty} \frac{dx}{(\gamma + x^2)^{1/(1-s)}} = \int_0^{+\infty} \frac{\gamma^{1/2} dx}{\gamma^{1/(1-s)} (1+x^2)^{1/(1-s)}} = \frac{\gamma^{1/2-1/(1-s)}}{2} B\left(\frac{1}{1-s} - \frac{1}{2}, \frac{1}{2}\right)$$

を得る。一方、$\int_0^{+\infty} \psi(x) dx$ が収束するので定理 2 の証明より、$\int_0^{+\infty} x^2 \psi(x) dx$ が収束するための必要十分条件は、$\int_0^{+\infty} \psi(x)^s dx$ が収束することである。

$$\int_0^{+\infty} \frac{dx}{(\gamma + x^2)^{s/(1-s)}} = \int_0^{+\infty} \frac{\gamma^{1/2} dx}{\gamma^{s/(1-s)} (1+x^2)^{s/(1-s)}} = \frac{\gamma^{1/2-s/(1-s)}}{2} B\left(\frac{s}{1-s} - \frac{1}{2}, \frac{1}{2}\right)$$

であるが、ここで $s/(1-s) - 1/2 > 0$、即ち $s > 1/3$ でなければならない。このとき、

$$\int_0^{+\infty} \frac{x^2 dx}{(\gamma + x^2)^{1/(1-s)}} = \int_0^{+\infty} \frac{\gamma^{3/2} x^2 dx}{\gamma^{1/(1-s)} (1+x^2)^{1/(1-s)}} = \frac{\gamma^{3/2-1/(1-s)}}{2} B\left(\frac{1}{1-s} - \frac{3}{2}, \frac{3}{2}\right)$$

なので

$$\int_0^{+\infty} x^2 \psi(x) dx = \frac{\gamma}{2} \cdot \frac{B(1/(1-s) - 3/2, 3/2)}{B(1/(1-s) - 1/2, 1/2)} = \frac{\gamma}{2} \cdot \frac{1-s}{3s-1}$$

定理 3　$1/3 < s < 1$ のとき (**) の最適解は、次で与えられる。

$$\psi(x) = \frac{1}{Z}(\gamma + x^2)^{1/(s-1)}, \qquad x \geq 0$$
$$Z = \gamma^{1/2-1/(1-s)} B\left(\frac{1}{1-s} - \frac{1}{2}, \frac{1}{2}\right), \quad \gamma = \sigma^2 \cdot \frac{3s-1}{1-s}$$

$s = 1/3$ のとき

$$\int_0^{+\infty} \frac{dx}{\sqrt{\gamma + x^2}} = \left[\log\left(x + \sqrt{\gamma + x^2}\right)\right]_0^{+\infty}$$

は発散する。またこのとき、

$$\int_0^{+\infty} \frac{x^2 dx}{(\gamma + x^2)^{3/2}} = \frac{1}{2}\left[x\sqrt{\gamma + x^2} - \gamma \log\left(x + \sqrt{\gamma + x^2}\right)\right]_0^{+\infty}$$

も発散する。

$0 < s \leq 1/3$ のとき問題 (**) を考え直してみよう。まず $0 < s < 1/3$ とする。このとき、$\alpha > 0$ に対して

$$\phi(x) = \frac{(x+\alpha)^{-1/s}}{Z}, \qquad x \geq 0$$

なる密度関数を考える。Z は正規化定数で

$$Z = \int_0^{+\infty} (x+\alpha)^{-1/s} dx = \frac{s}{s-1}\left[(x+\alpha)^{(s-1)/s}\right]_0^{+\infty} = \frac{s\alpha^{(s-1)/s}}{1-s}$$

である。このとき、

$$\int_0^{+\infty} x^2 \phi(x) dx = \frac{1}{Z}\int_0^{+\infty} \frac{x^2 dx}{(x+\alpha)^{1/s}} = \frac{1}{Z}\int_0^{+\infty} \frac{\alpha^3 x^2 dx}{\alpha^{1/s}(x+1)^{1/s}} = \frac{\alpha^{3-1/s}}{Z} B(3, 1/s - 3)$$

より

$$\int_0^{+\infty} x^2 \phi(x) dx = \frac{\alpha^2(1-s)}{s} B(3, 1/s - 3)$$

であるので、$\int_0^{+\infty} x^2 \phi(x) dx = \sigma^2/2$ となる $\alpha > 0$ を見つけることができる。一方

$$\int_0^{+\infty} \phi(x)^s dx = \frac{1}{Z^s}\int_0^{+\infty} \frac{dx}{x+\alpha}$$

なので、この積分は任意の $\alpha > 0$ に対して発散する。よって、$0 < s < 1/3$ のとき問題 (**) は非有界である。

次に $s = 1/3$ の場合について考える。このとき、任意の $\alpha, \epsilon > 0$ に対して

$$\phi(x) = \frac{(x+\alpha)^{-(3+\epsilon)}}{Z}$$

なる密度関数を考える。Z は正規化定数で

$$Z = \int_0^{+\infty} (x+\alpha)^{-(3+\epsilon)} dx = \frac{1}{-2-\epsilon}\left[(x+\alpha)^{-2-\epsilon}\right]_0^{+\infty} = \frac{\alpha^{-2-\epsilon}}{2+\epsilon}$$

である。このとき、

$$\int_0^{+\infty} x^2 \phi(x) dx = \frac{1}{Z}\int_0^{+\infty} \frac{x^2 dx}{(x+\alpha)^{3+\epsilon}} = \frac{1}{Z}\int_0^{+\infty} \frac{\alpha^3 x^2 dx}{\alpha^{3+\epsilon}(x+1)^{3+\epsilon}} = \frac{\alpha^{-\epsilon}}{Z} B(3, \epsilon)$$

より

$$\int_0^{+\infty} x^2 \phi(x) dx = \alpha^2 (2+\epsilon) B(3, \epsilon)$$

であるので、任意の $\epsilon > 0$ に対して $\int_0^{+\infty} x^2 \phi(x) dx = \sigma^2/2$ となる $\alpha > 0$ を見つけることができる。一方

$$\int_0^{+\infty} \phi(x)^{1/3} dx = \frac{1}{Z^{1/3}}\int_0^{+\infty} \frac{dx}{(x+\alpha)^{1+\epsilon/3}} = \frac{-3}{Z^{1/3}\epsilon}\left[(x+\alpha)^{-\epsilon/3}\right]_0^{+\infty} = \frac{3(2+\epsilon)^{1/3}}{\epsilon\alpha^{2(1+\epsilon)/3}}$$

である。$\int_0^{+\infty} x^2\phi(x)dx = \sigma^2/2$ の条件の下で、即ち $\alpha^2(2+\epsilon)B(3,\epsilon) = \sigma^2/2$ の条件の下で、$\epsilon \to 0$ とすると $B(3,\epsilon) \to +\infty$ であるので、$\alpha \to 0$ である。従って、このとき

$$\int_0^{+\infty} \phi(x)^{1/3}dx = \frac{3(2+\epsilon)^{1/3}}{\epsilon\alpha^{2(1+\epsilon)/3}} \to +\infty$$

となる。よって、$s = 1/3$ のときも問題 $(**)$ は非有界である。

定理 4　$0 < s \leq 1/3$ のとき $(**)$ の最適解は非有界である。

定理 5　$s > 1$ とする。$\alpha, \beta \in \mathbb{R}$ および $M > 0$ が存在して、任意の $0 \leq x \leq M$ に対して $\alpha+\beta x \geq 0$ かつ任意の $x > M$ に対して $\alpha + \beta x < 0$ であり、

$$\psi(x) = \begin{cases} \left(\alpha + \beta x^2\right)^{1/(s-1)}, & x \leq M \\ 0, & x > M \end{cases}$$

である ψ が $(**)$ の実行可能解となるならば、それは最適解である。

証明　ψ を $(**)$ の任意の実行可能解とする。このとき、

$$\begin{aligned} \int_0^{+\infty} \psi(x)^s dx &= \int_0^M \psi(x)^s dx = \int_0^M \left(\alpha + \beta x^2\right)^{s/(s-1)} dx \\ &= \int_0^M \left(\alpha + \beta x^2\right)(\alpha + \beta x)^{1/(s-1)} dx = \int_0^M \left(\alpha + \beta x^2\right)\psi(x)dx = \frac{\alpha}{2} + \frac{\beta\sigma^2}{2} \\ &= \int_0^{+\infty} \left(\alpha + \beta x^2\right)\phi(x)dx \leq \int_0^{+\infty} \psi(x)^{s-1}\phi(x)dx \end{aligned}$$

が成立するので、定理 1 (3) により

$$\int_0^{+\infty} \phi(x)^s dx \leq \int_0^{+\infty} \psi(x)^s dx$$

導かれ、ψ が $(**)$ が最適解であることが証明された。■

この定理において、$\alpha > 0$ かつ $\beta < 0$ でなければならない。そこで、ψ を

$$\psi(x) = \begin{cases} \dfrac{1}{Z}\left(\gamma - x^2\right)^{1/(s-1)}, & x \leq \sqrt{\gamma} \\ 0, & x > \sqrt{\gamma} \end{cases}$$

$$\gamma > 0, \quad Z = 2\int_0^{\sqrt{\gamma}} \left(\gamma - x^2\right)^{1/(s-1)} dx$$

のように書き換えることができる。ここでベータ関数の基本公式 (岩波全書「数学公式 1」p.220)

$$\int_0^1 x^\alpha \left(1 - x^\lambda\right)^\beta dx = \frac{1}{\lambda}B\left(\frac{\alpha+1}{\lambda}, \beta+1\right), \quad (\alpha, \beta > -1, \lambda > 0)$$

より

$$Z = \gamma^{1/(s-1)+1/2}B\left(\frac{1}{2}, \frac{s}{s-1}\right)$$

であり、また

$$\int_0^{\sqrt{\gamma}} x^2\left(\gamma - x^2\right)^{1/(s-1)} dx = \gamma^{1/(s-1)-1/2}B\left(\frac{3}{2}, \frac{s}{s-1}\right)$$

なので、

$$\int_0^{+\infty} x^2\psi(x)dx = \gamma \cdot \frac{s-1}{3s-1}$$

定理 6　$s > 1$ のとき (**) の最適解は、次で与えられる。

$$\psi(x) = \begin{cases} \dfrac{1}{Z}\left(\gamma - x^2\right)^{1/(s-1)}, & x \leq \sqrt{\gamma} \\ 0, & x > \sqrt{\gamma} \end{cases}$$

$$Z = \gamma^{1/(s-1)+1/2} B\left(\frac{1}{2}, \frac{s}{s-1}\right), \quad \gamma = \sigma^2 \cdot \frac{3s-1}{s-1}$$

参考文献

[1] G. Hardy, J. E. Littlewood & G. Pòlya, *Inequalities*, Cambridge University Press, Cambridge, 1952.

[2] D. Prato & C. Tsallis, "Nonextensive foundation of Levy distributions," *Phys. Rev. E*, vol. 60, pp. 2398-2401, 2000.

[3] D. Ruelle, *Statistical Mechanics*, W. A. Benjamin, Inc., London, 1969.

[4] H. Suyari, "The unique non self-referential q-canonical distribution and the physical temperature derived from the maximum entropy principle in Tsallis statistics," LANL e-print cond-mat/0502298.

[5] I. J. Taneja, "On generalized entropies and applications," *Lectures in Applied Mathematics and Informatics* (edited by L. M. Riccardi), pp.107-169, Manchester University Press, Manchester, 1998.

[6] C. Tsallis, R. S. Mendes, & A. R. Plastino, "The role of constraints within generalized nonextensive statistics," *Physica A*, vol. 261, pp.534-554, 1998.