

代表値の平均性と中心性

鈴木, 譲
九州大学大学院人間環境学研究院人間科学部門共生社会学講座

<https://doi.org/10.15017/4771870>

出版情報 : 人間科学共生社会学. 8, pp.1-11, 2018-02-14. 九州大学大学院人間環境学研究院
バージョン :
権利関係 :

代表値の平均性と中心性

鈴木 譲

要 旨

相加平均は平均値として、標準偏差は散布度として、記述統計においては最もよく使われる指標である。一般に標準偏差は、相加平均の周りの散布度を表す指標として説明される。しかしながら、現在の代表値の説明においては平均性と中心性という2つの性質が混同されている。平均値は平均性に関する指標、散布度は中心性に関する指標であり、この2つの指標は本来別個に論じられるべきである。この2つの性質が混同されているために、現在の説明では、標準偏差の代わりに平均偏差や3次以上の積率がなぜ用いられないのか、あるいは、相乗平均の周りの散布度はなぜ用いられないのか、といった根本的な疑問には適切に答えることができない。

本稿では平均性と中心性を別個に論じることにより、平均偏差の考え方では中心点が一般には一意的に定まらず、また、一意的に定まる場合もその値は中央値であって相加平均ではないこと、また、3次以上の積率の考え方では中心点が相加平均には一致しないこと、そして標準偏差と同じ手法では相乗平均の周りの散布度を定義することができないことを示す。

キーワード：相加平均、相乗平均、標準偏差、散布度、積率

1 はじめに

通常、記述統計において最初に説明される代表値は相加平均（算術平均）であり、その際、相加平均は分布の中心性（central tendency）を示す指標として説明される（Blalock 1979: 56; Bohrnstedt and Knoke 1988: 70; Guttman and Wilks 1965 = 1968: 58; 岡田1972: 5; 大澤2005: 19; 篠崎1994: 19; 竹村1997: 14; 内山2002: 8）。そしてこの中心値としての相加平均と各値との差を偏差とよび、一連の偏差自乗値の相加平均を計算した結果が分散、その正の平方根が標準偏差であり、分布の散布度（dispersion）を示す指標として説明される。

以上は統計学のほぼどの文献にでも見られる、いわばきわめて一般的な平均値と散布度の説明であるが、実は代表値の平均性（mean tendency）と中心性という2つの性質が混同されている。本来、相加平均は平均性にもとづく指標であり、中心性はまた別の性質である。ところがほとんどの統計学の文献では、相加平均が無条件に中心性をも有するという前提の下で論理が

展開され、散布度の指標が定義されている。

結果的には相加平均は中心性も併せ持つことが分かるので、このように分散、標準偏差を定義しても、散布度の指標として論理的問題があるわけではない。しかしながら、初めから相加平均をあたかも中心性の指標のごとく定義し、それを用いて散布度の指標を構成する論理構成は決して好ましいものではない。

このような説明では、以下にあげる3つの根本的な疑問に答えることができない。第一は、平均偏差 (mean deviation, average deviation) との指標としての優劣である。散布度の指標として、一連の偏差の相加平均を用いることが適切でない理由は、言うまでもなく偏差の合計は必ず0になるからである。そこで偏差に対して何らかの変換が必要になるが、偏差の自乗を考えるよりも、偏差の絶対値を考える方が自然である。一連の偏差絶対値の相加平均を計算した結果を平均偏差とよぶが、なぜ標準偏差の方が平均偏差よりも好ましいかを説明する必要がある。

多くの文献であげられている理由は、数式で絶対値を扱う煩雑さである (Blalock 1979: 78; 西平 1985: 76-77; 岡田 1972: 9)。また篠崎 (1994: 27) や松原 (1996: 87) では平均偏差が余り用いられないことは述べられているが、理由については書かれていない。佐々木 (1999: 27) や内山 (2002: 8) では平均偏差について述べられてはいるが、標準偏差との対比については書かれていない。しかし、絶対値の数式上の煩雑さは、平均偏差と標準偏差を散布度の指標として比較した場合の原理的な優劣とは直接関係はない。

また Bohrnstedt and Knoke (1988: 80-81) は、平均偏差を計算する場合に、相加平均よりも中央値を用いて計算した方が値が小さくなることを理由としてあげている。確かに、後に述べるように平均偏差の考え方では、相加平均よりも中央値の方が中心性が高くなる。しかしながら、Bohrnstedt and Knoke (1988) では平均性と中心性が区別されていないために、この説明だけではなぜ平均偏差が散布度の指標として適切でないのが明確ではない。

第二の疑問は、仮に平均偏差を使わないとしても、なぜ3次以上の積率 (モーメント) を用いないのかである。分散では偏差の自乗総和 (偏差平方和) を用いて散布度を計算する。分散は偏差平方値の平均値であるが、一般に k を自然数として、偏差の k 乗総和を標本数で除した値、つまり偏差 k 乗値の平均値を、平均値の周りの k 次の積率とよんでいる。

松原 (1996: 87) は、偏差を自乗することによってデータの散らばり具合が、平均偏差よりも際立つと述べている。このように考えると、平均値、平均偏差が同じ2つの分布に対しても、標準偏差を比較することによって散布度を比較できるように思える。そのようなデータを具体的に構成してみる。まず標本数 n が2の場合には、平均偏差と標準偏差は必ず一致し、平均値と標準偏差が与えられれば、2つの値は「平均値 ± 標準偏差」として一意的に決定されるから、そのようなデータを構成することはできない。そこで n が3の場合を考え、2つの集合 A と B を次のように定義する。(構成の仕方は Appendix 1 を参照)

$$A = \{0, 2, 4\} \quad B = \{0, 3, 3\}$$

A と B はともに平均値は2、平均偏差は $4/3$ であるから、これだけでは散布度の大小は比較

できない。しかし、標準偏差の値は A は 1.63、 B は 1.41 であるから、 A の方が散布度が大きいことになる。以上をまとめると表 1 の通りである。

表 1 相加平均と平均偏差が等しい分布

	相加平均	平均偏差	標準偏差
A	2	$4/3$	1.63
B	2	$4/3$	1.41

この説明は一見分かりやすいが、このように考えると、平均値と標準偏差が同じ 2 つの分布で、より高次の積率を用いないと散布度を比較できないような場合はないかが疑問として残る。そこで、そのようなデータを具体的に構成してみる。同じく n が 3 の場合を考え、集合 C を次のように定義する。(構成の仕方は Appendix 2 を参照)

$$C = \left\{ \frac{1}{2}, \frac{11 + \sqrt{21}}{4}, \frac{11 - \sqrt{21}}{4} \right\}$$

2 つの集合 B と C を比べると、ともに平均値は 2、分散は 2 である。そこで、 k を偶数として k 次の積率を比較してみる。4 次の積率はともに 6 で等しい。6 次の積率は B が 22、 C は 19.27 で B の方が大きい。従って、分散と 4 次の積率では分からないが、6 次の積率を比べると B の方が散布度が大きいことになる。以上をまとめると、表 2 の通りである。

表 2 相加平均と標準偏差が等しい分布

	相加平均	分散	4 次積率	6 次積率
B	2	2	6	22.00
C	2	2	6	19.27

しかし、このように高次の積率を用いるのでは、次数をいくらでも大きくする必要があり際限がなく、結局散布度の指標を適切に構成することができなくなってしまふ。

第三の疑問は、なぜ相乗平均（幾何平均）を用いないかである。周知の通り、平均値には相加平均、相乗平均、調和平均がある。平均値を中心性の指標として用いる、という場合は暗黙に相加平均を想定しており、他の平均に関しては考慮されていない。しかし、相乗平均も平均値の一種である以上、なぜ相加平均を用いて相乗平均は用いないのか、つまり、なぜ相乗平均の周りのデータのばらつきを考えないのかを説明する必要がある。

以上の 3 つの疑問は、平均性と中心性を区別して論理を展開しない限り適切に答えることができない。以下、この 2 つの性質を区別して論じ、これらの疑問に対する回答を提示する。

2 平均性の指標

「平均」とは言うまでもなく「平（たいら）に均（なら）す」こと、すなわち「値の大小の凸

凹を不揃いでないようにすること」である。今、 n を自然数、 $1 \leq i \leq n$ とし、 x_i を実数とする。任意の実数 y に関して、各 x_i の値を y に揃えることを考える。この時、どのような演算にもとづいて値を揃えるかが問題であるが、演算としてはまず加算を考える。すなわち x_i が y より大きければその差を余剰分とし、 x_i が y より小さければその差を不足分とし、余剰分と不足分を相殺して最終的に過不足を見るわけである。

相殺した結果を y の関数として $m(y)$ と表すと次の通りである。

$$m(y) = \sum_{i=1}^n (x_i - y)$$

この値が正ならば余剰があり、負ならば不足があるわけである。この値の絶対値は、 x_i を y にならそうとしたときの不揃いの程度を示している。そこで、あらためて $m(y)$ を次のように定義し、集合 $\{x_i\}$ の加算にもとづく y の不均整指数 (irregularity index) とよぶことにする。この値が小さいほど不揃いの程度が小さく、平均性が大きいといえる。この値が最小、つまり0となる場合が平均性が最大となる場合であり、その時の y の値が相加平均である。

$$m(y) = \left| \sum_{i=1}^n (x_i - y) \right|$$

調和平均は、逆数を用いた形での相加平均の応用である。 x_i を正の実数、 $t_i = 1/x_i$, $u = 1/y$ として、集合 $\{t_i\}$ の加算にもとづく u の不均整指数を考え、平均性が最大となる u の値を求めれば、その時の y の値が $\{x_i\}$ の調和平均である。

次に、乗算を演算として用いて値を揃えるのが、相乗平均である。すなわち、 x_i と y を正の実数とする時、 x_i が y より大きければその比率を余剰分とし、 x_i が y より小さければその比率を不足分とし、余剰分と不足分を相殺して最終的に過不足を見るわけである。乗算を演算としているので、過不足の見方は各比率の総積 (総乗) によって行い、その結果が1より大きければ余剰があり、1より小さければ不足があるわけである。

そこで、 $m(y)$ を次のように定義し、集合 $\{x_i\}$ の乗算にもとづく y の不均整指数とよぶことにする。

$$m(y) = \left| 1 - \prod_{i=1}^n \frac{x_i}{y} \right|$$

相加平均の場合と同様に、この値が小さいほど不揃いの程度が小さく、平均性が大きいといえる。この値が最小、つまり0となる場合が平均性が最大となる場合であり、その時の y の値が相乗平均である。

このように、平均値は相加平均、相乗平均のいずれにおいても、あくまで平均性にもとづいて定義される代表値であり、中心性にもとづいて定義されるわけではない。

文献によっては、物理的な重心との関係で、平均を説明しているものもある (岡田 1972: 6)。確かに、ベクトルとしての平均は物理学での重心と一致する。しかしながら重心としての説明、あるいは、ベクトルの平均としての説明は、相加平均には適用できるが、相乗平均に関しては

適用できない。従って、この方法では平均性を相加平均、相乗平均に共通の性質として説明することはできない。また、重心という用語は「中心」という意味を含んでいるが、これはあくまで物理学における質点に関する中心であって、以下に述べる散布度における中心性ではない。

3 中心性の指標

分布の散布度を論じる場合、どの点の周りにデータが散らばっているかが問題であるから、まず分布の中心点を定義する必要がある。今、 n を自然数、 $1 \leq i \leq n$ とし、 x_i を実数とする。任意の実数 y に関して、 y の周りに $\{x_i\}$ が散らばっている程度を示す関数 $f(y) \geq 0$ を考え、散乱指数 (scattering index) とよぶことにする。散乱指数の値が小さいほど y の周りのデータの散らばりが小さく、 y の持つ中心性が大きいと考えられる。そこで f が最小値をとるような y が一意的に存在するとき、この y の値を集合 $\{x_i\}$ の f にもとづく中心点とよび、その y に対する散乱指数の値を、集合 $\{x_i\}$ の f にもとづく散布度とよぶことにする。

円の中心などを想定すると、中心点としては、各値から等距離にある点、つまり各値からの偏差が等しくなる点が望ましいように思えるが、これは正しくない。もちろん、各値からの偏差が等しくなるような点は一般には存在しないが、たとえ存在したとしても、それは散布度を考えるための中心点の概念とは一致しない。

たとえば、 a と b を実数とし $a < b$ とする。ここで、 $\{a, a, \dots, a, b, b, \dots, b\}$ という集合を考えると a, b それぞれの要素数には関係なく、 $(a+b)/2$ はすべての値から等距離にある。この方法では異なる値が何種類あるかだけが問題となる。今異なる値は2種類しかない以上、中心点は常に $(a+b)/2$ であり、散布度は常に $(b-a)/2$ となる。しかし、散布度を考える場合には、たとえ同じ値であってもそれが複数ある場合には、それなりの重みづけをして考える必要がある。

従って、各値からの偏差が等しいかどうかは $\{x_i\}$ における y の中心性の判断基準として適切ではない。散乱指数として、 y の周りでの $\{x_i\}$ のばらつきの程度を測るためには、偏差を何らかの形で統合する必要がある。

散乱指数として偏差平方の平均値を考えると、 f は以下のように定義される。

$$f(y) = \frac{1}{n} \sum_{i=1}^n (x_i - y)^2$$

周知の通り、 f が最小値をとる y は一意的に定まり、その値は相加平均に等しい。従って、相加平均は偏差平方の平均値にもとづく中心点である。結果として、相加平均は平均性に加えて、中心性も併せ持つことが分かる。散布度としては上記の散乱指数の値、つまり分散を用いるわけである。

4 平均偏差

次に散乱指数として偏差絶対値の平均値を考えると、 f は以下のように定義される。

$$f(y) = \frac{1}{n} \sum_{i=1}^n |x_i - y|$$

x_i は昇順に並べ替えているものとする。 n が奇数であるか、偶数で $x_{n/2} = x_{(n/2)+1}$ であれば、 f は中央値（メジアン）で最小値をとるが、 n が偶数で $x_{n/2} \neq x_{(n/2)+1}$ であれば、 f が最小値をとる y の値は一意には定まらず、この2値の間の任意の値で最小値をとる。つまり f にもとづく中心点は一点には定まらない。

従って、この散乱指数を用いて散布度を測れるのは、 n が奇数であるか、偶数で $x_{n/2} = x_{(n/2)+1}$ の場合に限られ、しかもその際の中心点は中央値である。言い換えれば、この方式は相加平均ではなく、中央値の周りのデータのばらつきを測っている。このように見ると、平均偏差とは、相加平均に対する上記の散乱指数の値であり、相加平均がこの散乱指数にもとづく中心点でない以上、散布度の指標としては適切ではないことが分かる。

5 高次の積率

次に k を自然数として、散乱指数として y の周りの k 次の積率を考えると、 f は以下のように定義される。通常、積率の計算では絶対値をつけないが、ここでは k が奇数の場合にも、ばらつきの程度を正の値として測れるように、絶対値をつけた形で定義している。もちろん、 k が偶数であれば絶対値をつけてもつけなくても同じである。 k が 1 と 2 の場合は平均偏差と分散に対応し、すでに述べたので、 k が 3 以上の場合を考える。

$$f(y) = \frac{1}{n} \sum_{i=1}^n |x_i - y|^k$$

ここで f は連続関数であり、 $\lim_{y \rightarrow +\infty} f(y) = \lim_{y \rightarrow -\infty} f(y) = \infty$ であるから、 f は必ず最小値を持つ。ただ一般に最小値をとる y の値は、相加平均には一致しない。（詳細は Appendix 3 を参照）従って、最小値をとるような y の値が一意的に定まるとしても、この散乱指数の中心点は相加平均にはならない。言い換えれば、この方式では中心としての相加平均の周りのデータのばらつきを測ることにはならない。

6 相乗平均

最後に相乗平均について考える。 x_i と y を正の実数とする。相乗平均の場合には、演算として乗算を基礎にしているから、偏差に相当する値は $1 - x_i/y$ であり、この値の絶対値の k 乗平均を散乱指数として考えれば良く、 f は以下のように定義される。

$$f(y) = \frac{1}{n} \sum_{i=1}^n \left| 1 - \frac{x_i}{y} \right|^k$$

ここで f は連続関数で、 $\lim_{y \rightarrow 0} f(y) = \infty$ 、 $\lim_{y \rightarrow +\infty} f(y) = 1$ である。また、 y を十分大きくとれば、 $0 < 1 - x_i/y < 1$ すなわち $f(y) < 1$ が成り立つ。つまり横軸を y 軸、縦軸を z 軸とすると、 $y \rightarrow \infty$ の時、 $z = 1$ が漸近線となり、かつ y を十分大きくとると、 f はグラフとしてはこの漸近線の下に位置している。以上の条件から、 f は必ず最小値を持つことが分かる。

しかし、一般に最小値をとる y の値は、相乗平均には一致しない。(詳細は Appendix 4 を参照) 従って、この方式では相乗平均は中心点にはならず、中心としての相乗平均の周りのデータのばらつきを測ることはならない。

7 まとめ

本稿では、代表値の平均性と中心性の2つの性質を区別して論じ、平均値と散布度との関係を明確にした。平均値に関しては不均整指数を、散布度に関しては散乱指数を定義し、これらの指数が最小になる場合に、平均性と中心性が最大になると考えた。このように見ると、最頻値(モード)には対応する散乱指数がないので、散布度を考える上での中心性は持っていないことが分かる。また、調和平均は不均整指数から直接は定義されないが、逆数を用いた相加平均の応用であるので、平均性を持っていると解釈できる。また中央値は、 n が奇数、もしくは、偶数で中央の二項が等しい場合に限っては、中心性を持つので、限定的な中心性の指標といえる。以上をまとめると表3の通りである。

表3 代表値の平均性と中心性

	平均性	中心性	散布度
相加平均	○	○	標準偏差
相乗平均	○	×	—
調和平均	○	×	—
中央値	×	△	偏差絶対値の平均
最頻値	×	×	—

表3から明らかなように、平均性と中心性をともに備えている代表値は相加平均のみである。これが、相加平均と分散(標準偏差)が指標として最もよく使われる論理的理由である。

本稿では離散分布のみを扱った。積率は連続分布においても定義されるが、平均偏差と相乗平均は連続分布においては定義されない。また、連続分布においては平均値、分散が存在するとは限らない。従って、平均性と中心性についての包括的な議論のためには、連続分布よりも離散分布の方が適していると考えられる。

Appendix 1 相加平均と平均偏差が同じ分布

集合 $A = \{0, 2, 4\}$ は、相加平均 2、平均偏差 $4/3$ を持つ。集合 $\{0, x, y\}$ が同じく相加平均 2、平均偏差 $4/3$ を持つように x と y を定める。この 2 条件は、以下のように記述できる。

$$x + y = 6$$

$$|x - 2| + |y - 2| = 2$$

この 2 条件を整理すると、 $x, y \geq 2$ で $x + y = 6$ であれば、常にこの 2 条件を満足することが分かる。ここで $x = y = 3$ とすると、集合 $B = \{0, 3, 3\}$ を得る。

Appendix 2 相加平均と標準偏差が同じ分布

集合 $\{a, x, y\}$ が相加平均 m 、分散 s^2 を持つとする。この時、以下の 2 式が成り立つ。

$$x + y = 3m - a$$

$$(a - m)^2 + (x - m)^2 + (y - m)^2 = 3s^2$$

この式を x について解くと、以下の通りである。

$$x = \frac{1}{2} \left((3m - a) \pm \sqrt{3(2s^2 - (a - m)^2)} \right)$$

集合 B は相加平均 2、分散 2 であるから、 $a = 1/2, m = 2, s^2 = 2$ とすると、 B と同じく相加平均 2、分散 2 であるような集合 C を得る。

$$C = \left\{ \frac{1}{2}, \frac{11 + \sqrt{21}}{4}, \frac{11 - \sqrt{21}}{4} \right\}$$

Appendix 3 3 次以上の積率の最小値

$n = 3$ として、集合 $\{0, 1, 5\}$ を考えると、この集合の相加平均は 2 である。 k を 3 以上の自然数とする時、 k 次の積率 $f(y)$ は必ず最小値を持つが、 $y = 2$ では f は最小値とはならない。以下、 k が偶数の場合と奇数の場合とに分けて、この点を説明する。

積率において係数 $1/n$ は定数であるから、これを省いた偏差の k 乗総和について議論すれば良い。すなわち次式を考えれば良い。

$$f(y) = \sum_{i=1}^n |x_i - y|^k$$

まず、 k が偶数の場合を考える。 f は次式に示すように、全域で微分可能な関数である。 f の 1 次と 2 次の導関数は以下の通りである。

$$f(y) = y^k + (y - 1)^k + (y - 5)^k$$

$$f'(y) = k(y^{k-1} + (y - 1)^{k-1} + (y - 5)^{k-1})$$

$$f''(y) = k(k - 1)(y^{k-2} + (y - 1)^{k-2} + (y - 5)^{k-2})$$

f の 2 次の導関数は常に正であるから下に凸である。従って f は全域で最小値を持つ。その最小値は極小値でもあるから、その時の導関数の値は 0 となる。しかし、 k が 4 以上の偶数である

から、 $k-1$ は 3 以上の奇数であり、 $y=2$ の時、整数論におけるフェルマーの最終定理により、 f の導関数は 0 にはならないことが分かる。すなわち、次式が成り立つ。

$$2^{k-1} + 1 \neq 3^{k-1}$$

次に、 k が奇数の場合を考えると、 f は次式で与えられる。

$$f(y) = |y|^k + |y-1|^k + |y-5|^k$$

$f(y)$ は式としては、 $(-\infty, 0]$ $[0, 1]$ $[1, 5]$ $[5, \infty)$ の 4 つの区間でそれぞれ異なるが、ここでは相加平均 $y=2$ が含まれている $[1, 5]$ の区間についてだけ考えれば良い。 f は次式に示す通り、この区間で微分可能な関数である。 f の 1 次と 2 次の導関数は以下の通りである。

$$f(y) = y^k + (y-1)^k + (5-y)^k$$

$$f'(y) = k(y^{k-1} + (y-1)^{k-1} - (5-y)^{k-1})$$

$$f''(y) = k(k-1)(y^{k-2} + (y-1)^{k-2} + (5-y)^{k-2})$$

2 次の導関数は常に正であるから下に凸であり、この区間での最小値を持つ。もし全域での最小値がこの区間にあれば、その値はこの区間の最小値と一致し、その値は極小値でもあるから、その時の導関数は 0 となる。 k が 3 以上の奇数であるから、 $k-1$ は 2 以上の偶数である。まず、 $k=3$ の時には、

$$f'(2) = 3(2^2 + 1^2 - 3^2) = -12 \neq 0$$

次に k が 5 以上の奇数であれば、 $k-1$ は 4 以上の偶数であり、 $y=2$ の時、 k が偶数の場合と同様に、整数論におけるフェルマーの最終定理により、 f の導関数は 0 にはならないことが分かる。すなわち、次式が成り立つ。

$$2^{k-1} + 1 \neq 3^{k-1}$$

従って、この集合に関しては、3 以上のいかなる k についても、積率 $f(y)$ は相加平均 $y=2$ では最小値をとらないことが分かる。

Appendix 4 相乗平均の散乱指数

積率の場合と同様に係数 $1/n$ は定数であるから、これを省いた k 乗総和について議論すれば良い。

$n=2$ として、集合 $\{1, 3\}$ を考えると、相乗平均は 3 の正の平方根である。 k を自然数として散乱指数は次式で与えられる。

$$f(y) = \left|1 - \frac{1}{y}\right|^k + \left|1 - \frac{3}{y}\right|^k$$

$f(y)$ は式としては、 $(-\infty, 1]$ $[1, 3]$ $[3, \infty)$ の 3 つの区間でそれぞれ異なるが、ここでは相乗平均 $y \approx 1.73$ が含まれている $[1, 3]$ の区間についてだけ考えれば良い。この区間で f は微分可能な関数である。 f の導関数は以下の通りである。

$$f(y) = \left(1 - \frac{1}{y}\right)^k + \left(\frac{3}{y} - 1\right)^k$$

$$f'(y) = \frac{k}{y^2} \left(\left(1 - \frac{1}{y}\right)^{k-1} - 3 \left(\frac{3}{y} - 1\right)^{k-1} \right)$$

f の全域での最小値がこの区間にあれば、その値は極小値でもあるから、導関数は0となる。まず、 $k=1$ の場合は、導関数は常に負であるから、 f は極値をとらない。 k を2以上として、この導関数が0になる値を求めると、次の通りである。

$$y = \frac{1 + 3^{\frac{k}{k-1}}}{1 + 3^{\frac{1}{k-1}}}$$

$k=2$ の時、 $y=2.5$ である。 y を k の関数と見て、 y の k に関する導関数を計算すると、次の通りである。

$$y'(k) = -\frac{2 \cdot (\log 3) \cdot 3^{\frac{1}{k-1}}}{(k-1)^2} < 0$$

従って、 y は k に関して単調減少である。また、次式が成り立つ。

$$\lim_{k \rightarrow \infty} y(k) = 2$$

今、横軸を k 軸、縦軸を z 軸とすると、 y のグラフは $z=2$ を漸近線とし、常にこの漸近線よりも上に位置している。従って、 y の値が3の正の平方根、つまり、この集合の相乗平均に等しくなることはない。言い換えれば、この集合においてはいかなる k に関しても、相乗平均の値では散乱指数は最小値にはならない。

文 献

- Blalock, Hubert M., Jr., 1979, *Social Statistics*, revised second edition, Boston, MA: McGraw-Hill.
- Bohrnstedt, George W., and David Knoke, 1988, *Statistics for Social Data Analysis*, second edition, Itasca, IL: F. E. Peacock.
- Guttman, I., and S.S. Wilks, 1965, *Introductory Engineering Statistics*, New York: John Wiley. (= 1968, 石井恵一・堀素夫訳『工科系のための統計概論』培風館.)
- 松原望, 1996, 『わかりやすい統計学』丸善.
- 西平重喜, 1985, 『統計調査法 改訂版』培風館.
- 岡田泰栄, 1972, 『統計学概論』共立出版.
- 大澤秀雄, 2005, 『基礎から学ぶ統計学』梓出版社.
- 佐々木正文, 1999, 『社会科学系学生のための統計学』共立出版.

- 篠崎信雄, 1994, 『統計解析入門』サイエンス社.
- 竹村彰通, 1997, 『統計』共立出版.
- 内山敏典, 2002, 『計量分析のための統計解析技法』晃洋書房.