

多変数順位相関係数

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 : 教授

<https://doi.org/10.15017/4755916>

出版情報 : 人間科学共生社会学. 7, pp.1-10, 2016-09-16. 九州大学大学院人間環境学研究院
バージョン :
権利関係 :



多変数順位相関係数

鈴木 譲

要 旨

現在の順位相関係数は2変数に限定されており、2次のクロス表分析手法にとどまっている。社会調査のデータ分析において、個人の選好や賛成・反対の程度などは本来離散変数であるにもかかわらず、これらを連続変数として扱い多重回帰分析などの手法を用いるのは、離散変数を用いた多変量解析の手法が不十分だからである。この問題を解決するために、離散変数を用いた「多変数順位相関係数」(MRCC: Multivariate Rank Correlation Coefficient)を構築する。

具体的には、クロス表を用いるのではなく、調査データを各行が回答者、各列が変数に対応する行列として表し、2つの行、つまり、2人の回答者のデータ対の集合を考える。2つの行の対応する座標の大小関係に応じて、この集合を部分集合の直和に分解し、各部分集合の要素数を用いて係数を構築する。

キーワード：順位相関、多変数、離散変数

1 はじめに

意識調査などで得られた個人の選好や賛成・反対の程度は順序属性を持っているが、分析上の変数としては厳密に考えれば離散変数であり、連続変数ではない。しかしながら現実には、Pearsonの積率相関係数や多重回帰分析など、連続変数を前提とした計量分析の手法が通常用いられている。

このような状況は決して好ましくないが、そこから脱却できない主たる原因は、順序属性を持つ離散変数に関する多変量解析の手法が不十分だからと考えられる。2変数の場合には、クロス表分析の手法として何種類かの順位相関係数が提唱されているが、3変数以上では適切な指標が存在しない。

もちろん、質的データ、あるいは、カテゴリカル・データを分析する多変量解析の手法としては、数量化法3類、双対尺度法、最適尺度法、コレスポンデンス・アナリシス、等質性分析などがある(竹内1989:371)。これらはすべて等価な分析手法である。しかしながら、これらの手法は次に述べる2つの理由から相関係数や多重回帰分析に取って代わることが難しい。第

一の理由は、これらの手法が対象としているのは順序属性を持たないデータであり、変数が順序属性を持っていたとしても、分析の過程では順序属性は無視されるからである。第二の理由は、これらの手法はいわばデータの構造分析を目的としており、相関係数、決定係数、順位相関係数のように一つの数値でデータ全体の関連性を示すわけではないからである。

以上の状況を鑑み、多変数順位相関係数 (MRCC: Multivariate Rank Correlation Coefficient) を構築する。これが、本稿の目的である。現在の分析手法を、扱われる変数の性質と変数の個数にもとづいてまとめると、表 1 の通りである。

表 1 分析手法の分類

	2 変量の解析	多変量解析
連続変数	積率相関係数	多重回帰分析
離散変数	順位相関係数	なし

具体的な手順としては、既存の順位相関係数を修正して 2 変数から多変数に拡張するのではなく、多変数における順位相関とは本来どのように考えるべきか、という問から出発して、適切な指標を構築する。2 変数の場合に、新たに構築した多変数順位相関係数が既存の各種順位相関係数とどのような関係にあるかは、その後で考察を行う。

2 多変数における順位相関の考え方

2.1 同順位を許容しない場合

$k (k \geq 2)$ を変数の個数とする。各変数 x_i は順位属性を持つものとする。従って、各変数の取る値は自然数と考えて構わない。 N を自然数の集合とすると、各サンプル $z = (x_i)_i$ は、 k 次元空間 N^k の点と考えられる。 n をサンプル数とし、2 つのサンプル z_1, z_2 を取り出し、サンプル対 $t = (z_1, z_2)$ を構成する。このような対の総数は、 n 個から 2 個を取り出す組み合わせの総数 ${}_nC_2$ である。この ${}_nC_2$ 個のサンプル対から成る集合 T を全体集合として、各サンプル対において各座標の値の大小比較を行い、その結果に応じて T を複数の部分集合に排他的に分類し、多変数における順位相関を分析する。

根本的な問題は、多変数の場合の「順位相関」をどのように定義するかである。すなわち、多変数の場合に相関の正、負、0 (無相関) を概念的にどのようにとらえるかである。本稿では順位属性を持つ変数だけを取り上げるので、以下では特に混乱を招かない限り、順位相関の代わりに単に相関という用語を用いる。

まず、多変数における相関とは「すべての変数が同時に規則的に変化すること」と定義する。ここでは「すべての変数」としているのので、結果的に以下に示すように同順位を持つ対は考察から除外することになる。ただ、このままでは定義としては曖昧なので、明確にすべき点がある。

まず「規則的に変化する」ということの意味であるが、これは各変数の変化を個別に見た場

合に、その変数の変化の方向が一貫していることと考える。各変数の変化の方向の組み合わせは、方向をすべて反転した場合は重複しているからこれを除いて考えると、全体で 2^{k-1} だけある。従って規則的な変化といっても、 2^{k-1} の規則性がある。言い換えれば、変化の方向の一貫性といっても 2^{k-1} だけの一貫性があることになる。つまり 2^{k-1} の異なった相関が共存しているわけである。

ここで、各変数を個別に見た場合に、その変数の変化の方向に一貫性があるかどうかは、1つのサンプル対ではただか1つの変化しか分からないから判断ができない。一貫性があるかどうかを見るためには、サンプル対が2つ以上必要であり、加えて、各サンプル対でその変数の値が変化している必要がある。当然、すべての変数について変化の一貫性を調べるためには、すべての変数において変数の値が変化している必要があるから、同順位を1つでも持つサンプル対は具体的な規則性を調べる対象にはならないことになる。

ただ、同順位を持つ対については2通りの場合を区別しなくてはならない。まず、同順位でない変数が1つ以上あるようなサンプル対である。この場合は、変化していない変数があるのに、変化している変数もある、という状況であるから、この対自体が、変数全体としての相関性はない、つまり無相関（相関0）を示していると考えられる。次に、すべてが同順位の場合である。すべてが同順位の場合には、変数全体として相関があって値が無変化なのか、それとも、無相関であるがたまたま値が無変化なのかは区別がつかない。無相関であれば、いくつかの変数の値が無変化だとしても、残りの変数については値が無変化の場合もあれば変化する場合もあるからである。従ってこのような対は、変数全体としての相関、無相関に関して、何ら有益な情報を提供していないと考えられる。

以上の議論から、全体集合 T は以下の部分集合の直和として表現される。言い換えれば、以下の分類カテゴリーに排他的に分類される。ここで、 P_i はそれぞれ異なる相関に対応している。

P_i 同順位が1つもない場合 ($1 \leq i \leq 2^{k-1}$)

R_i 同順位が1つ以上あり、同順位でない変数も1つ以上ある場合

($1 \leq i \leq (3^k - 2^k - 1)/2$)

R_0 同順位でない変数が1つもない場合（すべて同順位の場合）

R_0 は分類カテゴリーとしては単独であるが、多変数の相関、無相関に関して有益な情報を提供していないと考えられるから、考察から除外し、 T から R_0 の要素を除いた集合を S として、 S を全体集合としてとらえれば良い。あるいは、 S における条件付き確率を考えるととっても良い。

なお、 $0 \leq i \leq k-1$ とする時、 i 個の同順位の変数を持つような分類カテゴリーの個数は、次に示すように ${}_k C_i \cdot 2^{k-i-1}$ である。まず k 個の変数の内、どの変数に i 個の同順位を持つ変数を割り当てるとかは ${}_k C_i$ だけの組み合わせがある。そしてそれぞれの場合に、残りの $k-i$ 個の変数は同順位を持たないから、これら $k-i$ 個の変数を2通りの大小関係により分類する方法は全部で 2^{k-i} だけある。ただ、大小関係を反転させた場合は同じ分類カテゴリーであるから、この重複を取

り除くために 2^{k-i} を 2 で除した値、つまり 2^{k-i-1} を先に計算した ${}_kC_i$ に乗じれば良い。これとは別に、 $i=k$ の場合、つまり、すべて同順位の場合は R_0 に相当し、その分類カテゴリーの個数は 1 である。

同順位が 1 つもない場合は、 $i=0$ であるから、分類カテゴリーの個数は 2^{k-1} である。同順位が 1 つ以上あり、同順位でない変数も 1 つ以上ある場合は、 $i=1$ から $k-1$ までの和となるから、二項定理を用いて計算すると、 $(3^k-2^k-1)/2$ となる。すべての分類カテゴリーの総数は、 $i=0$ から k までの和で、やはり二項定理を用いて計算すると、 $(3^k+1)/2$ となる。

S が R_i と P_i の直和として表現されることは、無相関を示す対と、異なる相関を示す対の集合に排他的に分類されていることに相当する。各 P_i においては、変数の変化の規則性は一貫しており相関が認められるわけであるが、これはいわば局所的な規則性、相関であり、局所相関と呼ぶことができる。問題は S 全体として、いわば全域的、大域的にどのように相関を判断するかである。異なった局所相関が共存しているということは、いくつかの変数が同じように変化した場合、残りの変数の変化に関して、異なった規則性が混在していることを意味している。

無相関には、いくつかの変数が変化していないのに残りの変数が変化する場合と、いくつかの変数が同じように変化しているのに残りの変数の変化に規則性がない場合の 2 通りがある。前者は 1 つのサンプル対から判断できるが、後者は複数のサンプル対を用いて判断しなくてはならない。前者は R_i に対応し、後者が今回の判断に対応する。

局所的な規則性が異なった形で混在しているということは、全体として見れば規則性がない、つまり、無相関ということになる。従って大域的に相関をとらえるための考え方としては、異なった局所相関の組み合わせを無相関として相殺し、結果としてどのような局所相関が残るのかを判断することになる。ここで問題は、相殺のためにどのような組み合わせを用いれば良いかを判断する基準がないことである。異なった局所相関は、どのような組み合わせを用いても、いくつかの変数の変化は同じで、残りの変数の変化の規則性が必ず異なっている。つまり、異なった局所相関は、どのような組み合わせを用いても何らかの形で相殺可能である。しかし、それではどのようにして異なった局所相関を相殺して、結果的にどの局所相関が残るのかを判断することができない。

つまり、どの局所相関に対して他の局所相関を相殺するかという基準は、分析者が指定する必要がある。順序属性は値を反転させれば、変数の値の増減は入れ替わるから、順序属性の値を適切に設定して、すべての変数の変化が正となるような局所相関を基準として設定すれば良い。これは、座標軸を反転させると考えても同じである。このようにすれば、サンプル対の変数値の変化で正負の値が混在しているような局所相関は、すべて負相関として扱い、基準となる局所相関と相殺されると考えることができる。

以上の議論に基づき、基準となる局所相関を正相関とし、それ以外の局所相関を負相関として扱う。正相関のカテゴリーの対には 1 を付与し、負相関のカテゴリーの対には負の値を付与し、無相関のカテゴリーの対には 0 を付与して、重み付け平均、つまり、期待値を計算すれば

良い。ここで、負相関のカテゴリーの対に一律に -1 を付与する方法も考えられるが、ここでは座標軸で基準相関のカテゴリー、つまり象限に近いものと遠いものとで付与する値を区別することにする。その理由は、基準となる座標軸は分析者が決めたもので固定されており、座標軸を自由に入れ替えられるわけではないから、これを基準にして象限の遠近を考えるのは意味があると判断できるからである。

基準となる正相関のカテゴリーを $(+, + \dots +)$ と表す。これは、 $(-, - \dots -)$ と同値であるから、正相関と負相関で符号が異なるのは、最大でも、 k が偶数ならば $k/2$ 、 k が奇数ならば $(k-1)/2$ である。符号の異なる座標が少ないほど、局所相関としてのカテゴリーの類似性が高く相殺効果は小さいと考え、逆に符号の異なる座標が多いほど局所相関としてのカテゴリーの類似性が薄く相殺効果は大きい、と判断する。具体的には、異なる符号の個数を u とする時、 k が偶数であれば $-2u/k$ を各対に付与し、 k が奇数であれば $-2u/(k-1)$ を各対に付与する。この値は、半分の符号が異なる場合に -1 となり、1つしか符号が異なる場合には $-2/k$ または $-2/(k-1)$ となる。

異なる符号を u 個持つ相関カテゴリーの個数は ${}_k C_u$ である。ただし、 k が偶数で $u = k/2$ の場合には、相関カテゴリーの個数は ${}_k C_u / 2$ となる。これらのカテゴリーの要素数の合計を Q_u とする。また、正相関のカテゴリーの要素数を P 、無相関のカテゴリー R_i の要素総数を R とする。 k が偶数か奇数かに応じて、多変数順位相関係数 ω を次のように定義すれば良い。

$$\omega = \frac{P - \frac{2}{k} \sum u Q_u}{P + R + \sum Q_u}$$

$$\omega = \frac{P - \frac{2}{k-1} \sum u Q_u}{P + R + \sum Q_u}$$

これは、正相関のカテゴリー要素に 1、無相関のカテゴリー要素に 0、負相関のカテゴリー要素に $-2u/k$ または $-2u/(k-1)$ を対応させる確率変数の期待値である。 ω は -1 から 1 までの間の値を取るが $0, 1, -1$ の値を取るのは次の場合である。

$$\begin{aligned} \omega = 0: & \quad P = \frac{2}{k} \sum u Q_u \text{ または } P = \frac{2}{k-1} \sum u Q_u \\ \omega = +1: & \quad R = 0, Q_u = 0 \\ \omega = -1: & \quad R = 0, P = 0, u = k/2 \text{ または } u = (k-1)/2 \end{aligned}$$

2.2 同順位を許容する場合

以上の設定では、同順位を持つ対は無相関として扱ってきた。次に条件を緩和して、同順位を持つ対も、相関カテゴリーに含めることを考えてみる。まず、すべてが同順位の場合は、これまでと同様に考察から除外する。また、同順位でない変数が 1 つだけの場合は、複数の変数

間の相関を示しているとは言えないので、無相関として扱う。従って、許容できるのは、少なくとも2つは同順位でない変数がある場合である。しかしながら、同順位でない変数のためだけに新たな相関カテゴリーを作成すると、余りに相関カテゴリーの個数が増え分析が難しくなるので、既存の相関カテゴリーに含めて考えるのが妥当と判断できる。ただ、同順位を持つ変数については大小関係が存在していないので、結局、同順位でない変数に関する大小関係が共通であるような相関カテゴリーのどれかに属するものとして扱うことになる。

具体的には、同順位でない変数について、すべて正の相関を示しているものは、正の相関カテゴリーに属する対として扱い、それ以外のものは負の相関カテゴリーに属している対として扱う。ただ、同順位でない変数の個数を v とする時、正の相関としては v/k の値を付与し、負の相関としては k が偶数か奇数かに応じて $-2u/k$ または $2u/(k-1)$ の値を付与する。

負の相関の場合、付与する値は、結果としては同順位を許容しない場合と同じになる。この理由は次の通りである。今、同順位でない変数の個数を v とする。つまり、 k 個の変数の内、 v 個の変数が同順位ではなく、その v 個の変数の内、 u 個の変数の符号が異なっているわけである。

仮に、 k と v が偶数とすると、同順位が全くない場合に符号が異なる変数の最大数は $k/2$ であり、実際に同順位ではない v 個の変数の内で符号が異なる変数の最大数は $v/2$ である。実際に符号が異なっている変数の個数は u であるから、負の相関の程度としては次の値を付与すれば良い。

$$-\frac{u}{v/2} \cdot \frac{v/2}{k/2} = -\frac{2u}{k}$$

もし、 k が奇数で v が偶数であれば、次の値を付与すれば良い。

$$-\frac{u}{v/2} \cdot \frac{v/2}{(k-1)/2} = -\frac{2u}{k-1}$$

v が奇数であっても、 $v/2$ が $(v-1)/2$ に置き換わり約分されるので、結果は全く同じである。従って、付与する値は同順位を許容しない場合と同じに結果になり、 v には影響されないことが分かる。まとめると、 k が偶数か奇数かに応じて、多変数順位相関係数 ω は次のように定義される。

$$\omega = \frac{\frac{1}{k} \sum v P_v - \frac{2}{k} \sum u Q_u}{R + \sum P_v + \sum Q_u}$$

$$\omega = \frac{\frac{1}{k} \sum v P_v - \frac{2}{k-1} \sum u Q_u}{R + \sum P_v + \sum Q_u}$$

ω はやはり -1 から 1 までの値を取るが $0, 1, -1$ の値を取るのは次の場合である。

$$\omega = 0: \quad \sum v P_v = 2 \sum u Q_u \text{ または } \frac{1}{k} \sum v P_v = \frac{2}{k-1} \sum u Q_u$$

$$\omega = +1: \quad R = 0, Q_u = 0, v = k$$

$\omega = -1$: $R = 0, P_y = 0, u = k/2$ または $u = (k-1)/2$

従って $\omega = 1$ となるのは、同順位が全くない場合に限られる。また $\omega = -1$ の場合に関しては、 k が偶数の場合は、同順位が 1 つでもあれば $u \leq k/2 - 1$ であるから $u = k/2$ とはならない。つまり k が偶数ならば $u = k/2$ となるのは、同順位が全くない場合だけである。 k が奇数ならば $u = (k-1)/2$ となるのは、同順位が全くないか 1 つだけの場合である。

3 2変数の場合の順位相関係数

$k = 2$ つまり 2 変数の場合には、同順位を持つ対は、2 変数ともに同順位か、それとも、一方の変数だけが同順位か、のいずれかである。前者は R_0 に対応するので、考察の対象にはならない。後者は、同順位を許容しない場合、許容する場合のいずれにしても、無相関としての扱いになる。従って、2 変数の場合には、同順位の許容に関しては、いずれの立場をとっても違いはない。逆の言い方をすれば、同順位を持つ対に関する多変数の場合の立場を明確にしておかないと、2 変数の場合をまず考えて、それを単純に 3 変数以上に拡張しようとしてもうまく行かない。

$k = 2, u = k/2 = 1$ であるから、異なる符号を u 個持つ相関カテゴリーの個数は ${}_2C_1/2 = 1$ である。このカテゴリーの要素数を Q 、正相関のカテゴリーの要素数を P 、無相関のカテゴリーの要素数を R とする。既存の文献の表記では、2 変数 x, y に関して、同順位がなく大小関係が揃っている対の個数を P 、同順位がなく大小関係が逆の対の個数を Q 、 x または y のいずれか一方だけが同順位の対の個数をそれぞれ X_0, Y_0 としているが、 P と Q はここでの定義と一致しており、 $R = X_0 + Y_0$ である。

多変数順序相関係数 ω は、定義に基づいて計算すると次の通りである。

$$\omega = \frac{P - \frac{2}{k} \sum u Q_u}{P + R + \sum Q_u} = \frac{P - Q}{P + Q + R} = \frac{P - Q}{P + Q + X_0 + Y_0}$$

ここでは、負の相関の対に一律に -1 を付与しても、異なる符号を持つ変数の個数の割合を用いても同じ結果となる。この式の分子は Kendall の τ_b もしくは Goodman-Kruskal の γ の定義と同じである。 τ_b もしくは γ との差異は R の扱いであり、この違いは上式の分母に反映されている。

他の順位相関係数としては、Somers の d があるが、この指標は従属変数と独立変数とを区別しているので、設定の前提そのものが異なり今回の比較の対象とはならない。また Stuart の τ_c は Kendall の τ_b の補正である。クロス表の行数と列数とが等しくない場合は必ず同順位が発生するから、 X_0, Y_0 のいずれかが 0 ではなく、結果として τ_b は 1 または -1 にはならない。Stuart の τ_c はこの点を補正したものである。しかしながら、Pearson の相関係数 r に関しても、 x と y の取る「異なる値」がそれぞれ何種類あるかを比較した場合、もしその数が等しくなければ回帰直線上にすべての点が当てはまることはなく、 r は 1 または -1 にはならない。これは

基本的にはクロス表で行数と列数とが等しくない場合に相当するから、このこと自体にそもそも大きな問題はないと判断できる。

従って、今回比較すべき対象は、Kendall の τ_b および Goodman-Kruskal の γ と考えられる。これらを列記すると次のようになる（安田・海野 1977: 42）。

$$\omega = \frac{P-Q}{P+Q+X_0+Y_0}$$

$$\tau_b = \frac{P-Q}{\sqrt{P+Q+X_0}\sqrt{P+Q+Y_0}}$$

$$\gamma = \frac{P-Q}{P+Q}$$

この比較から分かるように、今回定義した多変数順位相関係数 ω は、既存の順位相関係数 τ_b もしくは γ のいずれとも一致しない。まず τ_b では、分母での X_0 と Y_0 の扱い方の論拠が不明確である。Goodman and Kruskal (1954) は、この点を改善するために γ を提唱したと思われるが、 γ では X_0 と Y_0 が完全に無視されてしまっている。 ω はこれらの問題点を次のように解決している。 R は無相関の要素数を意味しているので、期待値を計算する場合に、分子には含めないが、分母には含めるべきである。具体的には、全体集合を構成する部分集合として、正相関の要素数と負相関の要素数に加算すべきである。

ω が 0, 1, -1 の値を取るのは次の場合である。

$$\omega = 0 : P = Q$$

$$\omega = +1 : Q = 0, X_0 = Y_0 = 0$$

$$\omega = -1 : P = 0, X_0 = Y_0 = 0$$

τ_b と ω で上記 3 つの条件は同一である。また、 γ については 0 となる条件は同一であるが、 γ が 1 または -1 になるのは、 $Q = 0$ または $P = 0$ の時であるから、 ω とは条件が異なる。

4 3変数の場合の順位相関係数

$k = 3$ つまり 3 変数の場合には、同順位の変数が 1 個で、同順位でない変数が 2 個の場合があるから、同順位の許容の仕方によって、順位相関係数の定義は異なる。また、負の相関に関しては、 k が奇数であるから $u = (k-1)/2 = 1$ で、符号が異なるのは最大でも 1 変数だけであることが分かる。異なる符号を u 個持つ相関カテゴリーの個数は ${}_3C_1 = 3$ である。

まず、同順位を許容しない場合は、正相関の要素数を P 、負相関の要素数を Q 、無相関の要素数を R とすると、多変数順位相関係数 ω は次式で与えられる。

$$\omega = \frac{P - \frac{2}{k-1} \sum u Q_u}{P + R + \sum Q_u} = \frac{P-Q}{P+R+Q}$$

この場合、負の相関の対に一律に -1 を付与しても、異なる符号を持つ変数の個数の割合を用いても同じ結果となる。 ω が $0, 1, -1$ の値を取るのは次の場合である。

$$\omega = 0 : P = Q$$

$$\omega = +1 : Q = 0, R = 0$$

$$\omega = -1 : P = 0, R = 0$$

次に、同順位を許容する場合は、同順位がない場合の正相関の要素数を P 、同順位がない場合の負相関の要素数を Q とし、同順位が 1 個の場合の正相関の要素数を P_1 、同順位が 1 個の場合の負相関の要素数を Q_1 とし、同順位が 2 個の場合の要素数の総数を R とすると、多変数順位相関係数 ω は次式で与えられる。この場合も、負の相関の対に一律に -1 を付与しても、異なる符号を持つ変数の個数の割合を用いても同じ結果となる。

$$\begin{aligned}\omega &= \frac{P + \frac{2}{3}P_1 - \frac{2Q}{3-1} - \frac{2Q_1}{3-1}}{P + P_1 + Q + Q_1 + R} \\ &= \frac{P + \frac{2}{3}P_1 - Q - Q_1}{P + P_1 + Q + Q_1 + R} = \frac{P - Q}{P + P_1 + Q + Q_1 + R} + \frac{\frac{2}{3}P_1 - Q_1}{P + P_1 + Q + Q_1 + R}\end{aligned}$$

同順位を許容しない場合と比較すると、第一項は同順位を許容しない場合の ω の値であるから、第二項が同順位を許容したことによる差を表している。差が正負のいずれになるかは、 P_1 と Q_1 の大小関係、つまり、同順位が 1 個の場合の正相関と負相関の要素数の大小関係に依存する。厳密には、同順位を許容した場合の方が多変数順位相関係数の値が大きくなる必要十分条件は、次の通りである。

$$P_1 > \frac{3}{2}Q_1$$

また、 ω が $0, 1, -1$ の値を取るのは次の場合である。

$$\omega = 0 : P + \frac{2}{3}P_1 = Q + Q_1$$

$$\omega = +1 : P_1 = 0, Q = 0, Q_1 = 0, R = 0$$

$$\omega = -1 : P = 0, P_1 = 0, R = 0$$

5 おわりに

本稿では、選好や賛否の程度を示す調査データは、厳密には連続変数ではなく、順序属性を持つ離散変数であるとの認識に立ち、より適切な多変量解析の指標を提唱した。なお連続変数を前提とする手法においても、2 変数の場合の相関係数は、3 変数以上には拡張されていない。

2 変数であれば Pearson の積率相関係数は 2 つの変数を同等に扱った形で関連性を示しているが、3 変数以上ではこのような指標は存在しないからである。多重回帰分析の決定係数は、1 つの数値で分析全体の様子を表してはいるが、従属変数と独立変数の設定に依存しており、全変数を対等に扱っているわけではない。ここで提唱した多変数順位相関係数はすべての変数を同等に扱っており、その意味で 2 変数の順位相関係数からの自然な拡張と言える。

文 献

Goodman, Leo A., and William H. Kruskal, 1954, “Measures of Association for Cross Classifications,” *Journal of the American Statistical Association*, 49: 732–764.

竹内啓編, 1989, 『統計学辞典』東洋経済新報社.

安田三郎・海野道郎, 1977, 『改訂2版 社会統計学』丸善株式会社.