1.5 順位相関係数

前節で取り扱った相関係数は，2つの対応する量と量との間の相関関係を尺度化したもので，ピアソンの相関係数とも呼ばれている．ところが実際の問題において，2つの変量をともに量的に表現することが困難で，単に順位だけなら測定しうる場合がある．このような場合，順位を用いて相関の度合いを尺度化することを考えなければならない．

スピアマンの順位相関係数

例 あるデパートで，A，B，C，D，E，F，Gの7種類の反物の柄，色について甲，乙2人に好きな順に順位をつけてもらったら次のようになった．

反物	A	B	C	D	E	F	G
甲の順位	2	4	1	7	3	5	6
乙の順位	5	4	7	3	6	2	1
差＝甲の順位－乙の順位	－3	0	－6	4	－3	3	5
差の2乗	9	0	36	16	9	9	25

この結果甲，乙2人の好みにどの程度の一致がみられるかを判断しよう．

一般に，反物の数が n のとき甲，乙の順位が次のようであったとする．

反物	A_1	A_2 ………	A_n
甲の順位	r_1	r_2 ………	r_n
乙の順位	s_1	s_2 ………	s_n

このとき甲と乙の順位の間の相関係数は式 (*1.14*) より次のようになる．

$$\rho=\frac{\sum_{i=1}^{n}(r_i-\bar{r})(s_i-\bar{s})}{\sqrt{\sum_{i=1}^{n}(r_i-\bar{r})^2\sum_{i=1}^{n}(s_i-\bar{s})^2}},\quad \text{ただし}\quad \bar{r}=\bar{s}=\frac{n+1}{2}$$

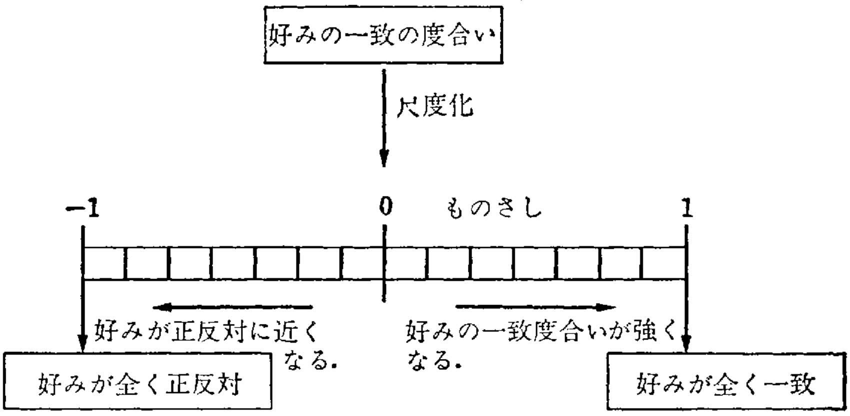
関係式：$\sum_{i=1}^{n}(r_i-\bar{r})(s_i-\bar{s})=\sum_{i=1}^{n}(r_i-\bar{r})^2-\frac{1}{2}\sum_{i=1}^{n}(r_i-s_i)^2,$

$$\sum_{i=1}^{n}(r_i-\bar{r})^2=\sum_{i=1}^{n}(s_i-\bar{s})^2=\frac{1}{12}(n^3-n)$$

を用いて上式を変形すると次式を得る．

$$(1.17)\qquad \rho = 1 - \frac{6\sum_{i=1}^{n}(r_i - s_i)^2}{n^3 - n}$$

これを**スピアマン**（Spearman）**の順位相関係数**と呼ぶ．好みの順位が全く同一ならば $\rho=1$，全く逆順位ならば $\rho=-1$ となることは相関係数の意味から容易にわかろう．したがって式（*1.17*）は次のように -1 から 1 までの長さをもつものさしで好みの一致の度合いを測っていることになる．



さて，甲，乙2人の7種類の反物の好みについては，式（*1.17*）から

$$\rho = 1 - \frac{6 \times 104}{7^3 - 7} = -0.86$$

となり，甲，乙2人の好みの一致性がみられないばかりか，むしろ好みは正反対に近いことがわかる．

また，次のような別の尺度化の方法も考えられているので紹介しよう．

ケンドールの順位相関係数

たとえば，反物の例において，7種類から2種類選ぶすべての組合せは次のようである．

(A，B)	(B，C)	(C，E)	(E，F)
(A，C)	(B，D)	(C，F)	(E，G)
(A，D)	(B，E)	(C，G)	(F，G)
(A，E)	(B，F)	(D，E)	
(A，F)	(B，G)	(D，F)	
(A，G)	(C，D)	(D，G)	

まず A, B の2つの組を考え，甲，乙2人の A, B についている順位の値を比較する．甲は A が2，B が4で，B の順位値の方が大きいが，乙は，A が5，B が4で，A の順位値の方が大きい．

甲: A の順位値＜B の順位値

乙: A の 〃 ＞B の 〃

のように順位の大小関係が反対になっているとき，A, B の組に対して点数 -1 を与え，A, E の組のように

甲: A の順位値＜E の順位値

乙: A の 〃 ＜E の 〃

のように順位の大小関係が一致するとき，A, E の組に対して点数 $+1$ を与える．このようにして，すべての組に点数を与える．

一般に，反物の数が n のとき，同様にして，全体の組の数 $\frac{1}{2}n(n-1)$ に点数を与え，次のような相関係数を考える．

$$(1.18)\quad \tau=\frac{K-L}{\frac{1}{2}n(n-1)}=\frac{2K}{\frac{1}{2}n(n-1)}-1=1-\frac{2L}{\frac{1}{2}n(n-1)}$$

ただし，K は点数 $+1$ の数，L は点数 -1 の数

$$\left(K+L=\frac{1}{2}n(n-1)\right)$$

これを**ケンドール** (Kendall) **の順位相関係数**，またはケンドールの τ (タウ) 係数と呼んでいる．好みの順位が全く同じなら $L=0$ となり $\tau=1$，好みの順位が全く逆順序なら $K=0$ となり $\tau=-1$ となることは式 (*1.18*) から容易にわかろう．

甲，乙2人の反物の好みについては，組合せと点数は次のようになる．

A, B -1	A, G -1	B, G -1	D, E -1	F, G -1
A, C -1	B, C -1	C, D -1	D, F $+1$	
A, D -1	B, D -1	C, E -1	D, G $+1$	
A, E $+1$	B, E -1	C, F -1	E, F -1	
A, F -1	B, F -1	C, G -1	E, G -1	

このとき，式 (*1.18*) から

$$\tau = \frac{3-18}{\frac{1}{2}\times 7\times 6} = -0.71$$

となり，スピアマンの順位相関係数の値 -0.86 とは異なっているが，-1 の方にかたよっているという同じ傾向を示し，好みは正反対に近いことがわかる．

以上，好みの一致性についてスピアマン，ケンドールの順位相関係数を述べたわけであるが，何も好みの一致性に限ったことはなく，これに類似するどのような問題にも適用できる．次節でも質問のところで，その1例を示している．

〈質問〉 スピアマン，ケンドールの順位相関のどちらを用いるかは，どのように使いわけしたらよいですか？

答　これは，どんな場合にスピアマンで，どんな場合にケンドールであるといった区別は特にありません．尺度化の方法が異なるだけで，どちらもほぼ同じ傾向を示します．たとえば，物の長さを測るのに単位が cm であるものさしを用いるか，単位が尺のものさしを用いるかといった区別と似ています．すなわち，順位相関係数の尺度化の意味をちゃんとわきまえていれば，どちらを用いても別にさしつかえありません．

問　　題　4

1. ある大学の学生食堂の日曜日から土曜日までの昼食の献立について，甲，乙2人の学生にその好みを聞いてみたら右表のようであった．

曜日	月	火	水	木	金	土
甲	3	5	1	6	2	4
乙	3	4	2	5	1	6

このとき，スピアマンの順位相関係数を求め，その結果この2人の学生の好みが一致しているかどうか判断せよ．

2. 1. においてケンドールの順位相関係数 τ を求めよ．

1.6　簡単な数値化の適用例

数値化の必要なことは，この章のはじめにも述べたが，最も簡単な数値化の例として，たとえば，次のように0と1の2つのタイプに分類するものがある．