第3回：2次元のデータAppendix

尚　晋
大学院経済学研究科　助教

2025年4月30日

ポイント

1. スピアマンの順位相関係数の証明について

1　スピアマンの順位相関係数の証明について

順位相関係数とは，二つの質的基準(量的変数を大小関係でこれに変換してもよい)がある場合に，観測対象iの，二つの基準による順位x_i, y_iの間の相関を示す指標である．

スピアマンの順位相関係数r_sは

$$r_s = 1 - \tfrac{6}{n^3-n}\Sigma_{i=1}^{n}(x_i - y_i)^2$$

と定義する．

スピアマンの順位相関係数の考え方は通常の積率相関係数の式を適用したものなので，まずは分母となるx, yのそれぞれの標準偏差を計算する必要があり，そして分子のC_{xy}共分散を計算します．

まずはxの分散を計算する．xは順位なので，平均値は$\bar{x} = \frac{n+1}{2}$，二乗の和は$\sum\limits_{i=1}^{n} x_i^2 = \frac{n(n+1)(2n+1)}{6}$（数列の二乗の和の公式）．順位なので，$\sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} y_i^2 = \sum\limits_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$も成り立つ．

従って，

$$\begin{aligned}
S_x^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \\
&= \frac{1}{n}(\Sigma x_i^2 - 2\bar{x}\Sigma x_i + n\bar{x}^2) \\
&= \frac{1}{n}(\Sigma x_i^2 - 2\bar{x}\cdot n\bar{x} + n\bar{x}^2) \qquad \Longleftarrow (\Sigma x_i = n\bar{x}\text{なので}) \\
&= \frac{1}{n}(\Sigma x_i^2 - 2n\bar{x}^2 + n\bar{x}^2) \\
&= \frac{1}{n}(\Sigma x_i^2 - n\bar{x}^2) \\
&= \frac{(n+1)(2n+1)}{6} - (\frac{n+1}{2})^2 \\
&= \frac{(n+1)(n-1)}{12}
\end{aligned}$$

次はyの分散も計算する．同様なので，結果は$S_y^2 = \frac{(n+1)(n-1)}{12}$.
また，

$$
\begin{aligned}
\sum_{i=1}^{n}(x_i - y_i)^2 &= \Sigma(x_i^2 - 2x_iy_i + y_i^2) \\
&= \Sigma x_i^2 - 2\Sigma x_iy_i + \Sigma y_i^2 \\
&= 2\Sigma i^2 - 2\Sigma x_iy_i \quad \Longleftarrow (\Sigma x_i^2 = \Sigma y_i^2 = \Sigma i^2 \text{なので})
\end{aligned}
$$

すなわち，書き換えると：

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

従って，共分散は：

$$
\begin{aligned}
C_{xy} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_iy_i - \bar{x}\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} i^2 - \frac{1}{2n}\sum_{i=1}^{n}(x_i - y_i)^2 - (\frac{n+1}{2})^2 \\
&= \frac{(n+1)(2n+1)}{6} - \frac{1}{2n}\sum_{i=1}^{n}(x_i - y_i)^2 - (\frac{n+1}{2})^2 \\
&= \frac{(n+1)(n-1)}{12} - \frac{1}{2n}\sum_{i=1}^{n}(x_i - y_i)^2
\end{aligned}
$$

これらの結果を全て積率相関関数の式に代入すると：
(積率相関関数の式は:$r_{xy} = \frac{C_{xy}}{\sqrt{S_x^2}\sqrt{S_y^2}}$で定義される．)

$$
\begin{aligned}
r_{xy} &= \frac{\frac{(n+1)(n-1)}{12} - \frac{1}{2n}\sum_{i=1}^{n}(x_i - y_i)^2}{\frac{(n+1)(n-1)}{12}} \\
&= 1 - \frac{6}{n^3 - n}\sum_{i=1}^{n}(x_i - y_i)^2
\end{aligned}
$$