



ケンドールの順序相関係数について

メタデータ	言語: Japanese 出版者: 北海道教育大学 公開日: 2012-11-07 キーワード: 作成者: 藤戸, 伊佐美 メールアドレス: 所属:
URL	https://doi.org/10.32150/00002679

ケンダールの順序相関係数について

藤 戸 伊 佐 美
北海道教育大学函館分校数学教室

On the Properties of the Kendall's Distribution

Isami FUJITO

Mathematics Laboratory, Hakodate College, Hokkaido University of Education,
Hokodate 040

Abstract

We consider a representation of the distribution which deduce Kendall's rank-difference correlation coefficient, and show several properties of distribution. Using these properties, we give an asymptotic behaviour of the distribution when the order of the distribution increases to infinity.

1 序

ケンダールの順序相関係数の母集団模型を構成し、相関係数の分散の公式を証明する。次いで、母集団模型の分布の性質を導く。最後に、この分布の性質を用いて、ケンダールの母集団は、対象個数の増加に伴い、正規母集団と異なるものに近づくことを示す。

2 ケンダールの母集団模型

X, Y を順序構造を持つ集合とする。 $X \times Y$ からの標本として、 $(x_i, y_i), i=1, 2, \dots, N$ を考える。 $i \neq j$ のとき、 $x_i \neq x_j, y_i \neq y_j$ とする。 $i \neq j$ のとき、 $\{(x_i, y_i), (x_j, y_j)\}$ に対して、 $x_i < x_j, y_i < y_j$ 又は、 $x_i > x_j, y_i > y_j$ のとき、 1 を対応させ、 $x_i < x_j, y_i > y_j$ 又は、 $x_i > x_j, y_i < y_j$ のとき、 -1 を対応させる。この対応で、 P を 1 をとる場合の数、 Q を -1 をとる場合の数とする。総数は、 ${}_N C_2 = T$ となるので、 $P + Q = T$ となる。 $P - Q = S$ とおくと、 S の分散 $V(S)$ は、

$$V(S) = \frac{N(N-1)(2N+5)}{18}$$

で与えられることは、周知のことである。

こゝで、 N 個の標本を固定し、 i を入れかえることにより、 $N!$ 個の標本の組を作る。この標本の組に対して、 S の分布が定まる。 $N=2$ のとき、 S の値と、その値をとる度数 f との関係は、

S	1	-1
f	1	1

$N=3$ のときは、

S	3	1	-1	-3
f	1	2	2	1

$N=4$ のときは、

S	6	4	2	0	2	-4	-6
f	1	3	5	6	5	3	1

となる。

$N=2$ から、 $N=3$ への分布の移行は、

	1	1		
		1	1	
			1	1
計	1	2	2	1

として、 $N=3$ から、 $N=4$ への分布の移行は、

	1	2	2	1		
		1	2	2	1	
			1	2	2	1
				1	2	2
計	1	3	5	6	5	3
						1

として、得られる。一般に、 $N-1$ から、 N への分布の移行については、 $N-1$ の分布を、右へ i 個平行移動したものを、 $i=0$ から、 $i=N-1$ まで作り、各列の和をとったものが、 N のときの分布となっている。

N のとき、 S のとり得る値は、

$$\frac{N(N-1)}{2} - K$$

で、 $K=0, 2, 4, \dots, N(N-1)$ である。従って、 S は、

$$\frac{N(N-1)}{2} + 1 = \frac{N^2 - N + 2}{2}$$

個の値をとる。 $N-1$ のときの S の値を、 x_1, x_2, \dots, x_m で、各 x_i をとる度数を f_i で示し、 N のときの S の値を、 y_1, y_2, \dots, y_n で、各 y_i をとる度数を F_i で示す。 m, n は、

$$m = \frac{N^2 - 3N + 4}{2}, \quad n = \frac{N^2 - N + 2}{2}$$

である。凡ての N について、得られた分布の平均は、0 となるから、 N のときの S の分散 $V_N(S)$ は、

$$V_N(S) = \frac{1}{N!} \sum_i F_i y_i^2$$

である。 $N-1$ から、 N への分布の移行過程より、

$$= \frac{1}{N!} \sum_i \sum_k f_i(x_i - a_k)^2$$

となる。ここで、 $a_k = (N-1) - 2k$, $k=0, 1, 2, \dots, N-1$ である。 $V_{N-1}(S)$ で、 $N-1$ のときの S の分散を表わすと、

$$= \frac{1}{N!} \left(\sum_k (N-1)! \cdot V_{N-1}(S) + (N-1)! \cdot \sum_k a_k^2 \right)$$

となる。

$$\begin{aligned} \sum_k a_k^2 &= \sum_{k=0}^{N-1} \{N - (2k+1)\}^2 \\ &= \sum_k (2k+1)^2 - N^3 \\ &= \frac{4N(N^2-1)}{3} + N - N^3 \\ &= \frac{1}{3} N(N^2-1) \end{aligned}$$

となるから、

$$V_N(S) = V_{N-1}(S) + \frac{1}{3}(N^2-1)$$

となる。帰納法を用いて、この関係式より、ケンドールの分散の公式が証明される。

3 ケンドール母集団の分布の性質

N のときの S の値を、 x_i で示す。 $i=1, 2, \dots, n$ で、

$$n = \frac{N^2 - N + 2}{2}$$

となる。 x_i をとる度数を、 $f_N(i)$ で示す。

$$\sum_i f_N(i) = N!$$

となる. x_i に, 確率 $f_N(i)/N!$ を与え, $i=1, 2, \dots, n$ としたときの全体を, ゲンドール母集団と呼ぶことにする. $j \leq N$ なる自然数 j に対して,

$$\sum_{i=1}^j f_{N-1}(i) = f_N(j)$$

となり, j を固定すると, $f_N(j)$ は, N についての $j-1$ 次の階差数列となる.

$$k_0 = \frac{(N-2)(N-3)}{2}$$

とおく. $k \leq k_0$ なる自然数 k に対して,

$$\sum_{i=1}^{k_0+k} f_{N-1}(i) = f_N(k_0+k)$$

となる. 又,

$$f_N(i) = f_N(n-i+1)$$

となる.

$$i_0 = \frac{N^2 - N + 4}{4}$$

とおく. i_0 が整数のとき, $f_N(i_0)$ は最大値となり, i_0 が整数でないとき, $f_N(i_0-1/2) = f_N(i_0+1/2)$ が最大値となる. f_N が, 最大値を2点でとると, f_{N-2} は, 最大値を1点でのみとり, f_N が, 最大値を1点でのみとるならば, f_{N-2} は, 最大値を2点でとる. いづれにしても, f_{N-2} の最大値をとる i の小さい方をとって, $[i]_{N-2}$ で示し, 又, f_N の最大値をとる i の大きい方をとって, $[i]_N$ で示すと $[i]_N - [i]_{N-2} = N-1$ となる. $f_{N-2}([i]_{N-2}) - f_{N-2}([i]_{N-2}-1) < f_{N-3}(i_1)$ となる. i_1 は, f_{N-3} の i についての定義域の中点より, $(N-2)/2$ ずれた i の値を示す. これが, 意味を持つためには, $N \geq 7$ を充たさねばならない. $f_{N-2}(i_1) > f_{N-3}(i_1)$ より, $f_{N-2}([i]_{N-2}) < f_{N-2}([i]_{N-2}-1) + f_{N-2}(i_1)$ となる. 従って,

$$\sum_{i < [i]_{N-2}} f_{N-1}(i) > \sum_{i < [i]_{N-2}} f_{N-2}(i) + \sum_{i < [i]_{N-2}-1} f_{N-2}(i) > \sum_{i \leq [i]_{N-2}} f_{N-2}(i) \geq \frac{(N-2)!}{2}$$

となる. 一方,

$$f_N([i]_N) = \sum_{i=[i]_{N-2}}^{[i]_{N-2}+(N-1)} f_{N-1}(i)$$

より,

$$\begin{aligned} f_N([i]_N) &< (N-1)! - (N-2)! \\ &= (N-2)! \cdot (N-2) \end{aligned}$$

となる．この不等式は， $N \geq 7$ のとき成立する．

他方，逆向きの不等式の評価式としては，

$$\left(\frac{(N-1)^2 - (N-1) + 2}{2} \right) \cdot \frac{1}{N} < \frac{N-2}{2}$$

が， $N \geq 5$ のとき，成立するから，

$$f_N([i]_N) \cdot N > N! \cdot \{(N-2)/2\}^{-1}$$

となり，

$$f_N([i]_N) > 2(N-1) \cdot (N-3)!$$

が，得られる．

次に，分布の増分について考える． $N-1$ から， N への分布の構成方法より，

$$f_N(i) = \sum_{j=1}^N f_{N-1}(i, j)$$

と書ける．ここで， $i \geq j$ で，

$$f_{N-1}(i, j) = f_{N-1}(i-j+1)$$

となる． $i < [i]_N$ のとき，

$$\begin{aligned} \Delta f_N(i) &= f_N(i) - f_N(i-1) \\ &= \sum_{j=1}^N (f_{N-1}(i, j) - f_{N-1}(i-1, j)) = f_{N-1}(i, 1) - f_{N-1}(i-1, N) < (N-3)! \cdot (N-3) \end{aligned}$$

となる．上の i の範囲で， $i = [i]_{N-1}$ のとき， $\Delta f_N(i)$ は最大値をとる．このことの証明をする．先づ， $[i]_N > i \geq [i]_{N-1}$ では，

$$\begin{aligned} f_{N-1}(i, 1) &> f_{N-1}(i+1, 1) \\ f_{N-1}(i-1, N) &< f_{N-1}(i, N) \end{aligned}$$

となるので，この範囲の i では， $\Delta f_N([i]_{N-1}) \geq \Delta f_N(i)$ となる．次に， $[i]_{N-1} > i \geq [i]_{N-2}$ では，

$$f_{N-1}(i+1, 1) - f_{N-1}(i, 1) = f_{N-2}(i+1, 1) - f_{N-2}(i, N-1) = f_{N-2}(i+1) - f_{N-2}(i+1-(N-1))$$

となる． f_{N-2} の， $i+1-(N-1)$ に於ける値は， $2[i]_{N-1} - 2[i]_{N-2} = N-1$ を考慮に入れると，

$$[i]_{N-2} + [i]_{N-2} - (i+1) + 2([i]_{N-1} - [i]_{N-2}) = 2[i]_{N-1} - (i+1)$$

に於ける値と等しい．更に，上式は， i の条件より，

$$> 2[i]_{N-1} - [i]_{N-1} = [i]_{N-1}$$

となるから，

$$f_{N-2}(i+1) - f_{N-2}(i+1-(N-1)) > 0,$$

となり、従って、

$$f_{N-1}(i+1, 1) - f_{N-1}(i, 1) > 0$$

となる。他方、

$$\begin{aligned} f_{N-1}(i, N) - f_{N-1}(i-1, N) &= f_{N-2}(i+N-1, 1) - f_{N-2}(i+N-2, N-1) \\ &= f_{N-2}(i+N-1) - f_{N-2}(i) < 0 \end{aligned}$$

となる。以上の結果より、 $[i]_{N-1} > i \geq [i]_{N-2}$ を満たす i でも、 $\Delta f_N(i)$ は、 $[i]_{N-1}$ でとる値を超えない。残りの、 $[i]_{N-2}$ 以下の i に対しては、 f_{N-2} の値が、単調に増加するので、 $\Delta f_N(i)$ も単調に増加する。従って、 $\Delta f_N(i)$ は、 $i = [i]_{N-1}$ と、 $i = N! - [i]_{N-1} + 1$ で最大値をとる。尚、 $\Delta f_N([i]_{N-1})$ の下からの評価式として、

$$\Delta f_N([i]_{N-1}) = f_N([i]_{N-1}) - f_N([i]_{N-1} - 1) = f_{N-1}([i]_{N-1}) - f_{N-1}([i]_{N-1} - N)$$

ここで、前に得られた評価式

$$f_{N-1}([i]_{N-1}) > 2(N-2) \cdot (N-4)!$$

と、

$$f_{N-1}([i]_{N-1} - N) < \frac{(N-3)!}{2}$$

とより、

$$\Delta f_N([i]_{N-1}) > \frac{(N-4)! \cdot (N-1)}{2}$$

が、得られる。

今まで、 S の値 x_i に対して、 $f_N(x_i)$ として表現すべき所を、 i の函数と考えて、 $f_N(i)$ として取り扱って来た。 f_N の定義域は、 $1/2 \cdot S$ を変数として取ると、 $N \equiv 0 \pmod{4}$ 、又は、 $N \equiv 1 \pmod{4}$ のときは、

$$\left[\frac{N(N-1)}{2}, -\frac{N(N-1)}{2} \right]$$

の区間内の整数となり、 $N \equiv 2 \pmod{4}$ 、又は、 $N \equiv 3 \pmod{4}$ のときは、

$$\left[\frac{N(N-1)}{2} + 1, -\frac{N(N-1)}{2} - 1 \right]$$

の区間内の整数となる。

4 $N \rightarrow \infty$ のときのケンドール母集団

$S' = aN^\alpha S$ と置く. a, α は定数とする. このとき, S' の分散 $V(S')$ は,

$$V(S') = a^2 N^{2\alpha} V(S) = \frac{a^2 N^{2\alpha} N(N-1)(2N+5)}{18}$$

となる. 従って, $a^2 = 9, \alpha = -3/2$ と置く, $N \rightarrow \infty$ のとき, $V(S') \rightarrow 1$ となる. 第3節の結果より, 分布の傾きが最大値をとる点は, N の1次式で表現されるから, 上の変換により, 傾きの最大値を示す点は, $N \rightarrow \infty$ とすると, 原点に近づく.

$$x_i - \frac{1}{2} N^{-\frac{3}{2}} < x < x_i + \frac{1}{2} N^{-\frac{3}{2}}$$

の x に対して,

$$\varphi_N(x) \cdot N^{-\frac{3}{2}} = \frac{f_N(x_i)}{N!}$$

とする. $\varphi_N(0)$ について,

$$\frac{2(N-1) \cdot (N-3)!}{N!} \cdot N^{\frac{3}{2}} < \varphi_N(0) < \frac{(N-2)! \cdot (N-2)}{N!} \cdot N^{\frac{3}{2}}$$

となる. 従って, $\varphi_N(0)$ は, $N \rightarrow \infty$ のとき, 高々, N の1/2乗程度で, $+\infty$ に近づくか, 又は, 一定値に近づくと思おぼせる.

φ_N の傾きの最大値は, 第3節の結果より,

$$\frac{(N-1) \cdot (N-4)!}{2 \cdot N!} \cdot N^3$$

以上の値を持ち,

$$\frac{(N-3) \cdot (N-3)!}{N!} \cdot N^3$$

以下の値を取る.

φ_N の原点での値が, N について有界であることが示されていない. 原点以外の他の点で, φ_N の値が N について非有界であるならば, 或る区間で, φ_N は N について非有界となるが, これは, 任意の N に対して, φ_N の作り方より,

$$\int \varphi_N dx = 1$$

であることに反する. 従って, 原点を含む任意の開区間を除けば, φ_N は一様収束する. この極限関数を φ と置く. φ は定義域を原点まで拡張出来るので, 定義域は, $(-\infty, +\infty)$ となる. 任意の N について, φ_N の分散が1であることより, φ の分散も1となることが示される. φ の原点での, 右側及び

左側微分係数も、 φ_N のそれらに対するものの極限として定義出来るので、上に得られた評価式よりこの2つの微分係数の絶対値は、等しくて、 $1/2$ 以上である。更に φ_N の原点に関しての対称性と、 φ_N の増分の最大値を取る点が原点に収束することと、 φ_N の増分の最大値を取る点までの φ_N の増分の単調性より、 φ の微分係数は、原点を除き単調に増加し、原点での片側微分係数の絶対値が、 ∞ を含めて、最大値となる。このことから、 φ は、原点で尖点を持つ分布函数となる。