相関係数の推定について

鍋　谷　清　治

相関係数の推定の精度については，これまでパラドックス的な結果が二三報告されているが，ここではそれらを含めて，いろいろな形の推定値の精度について調べてみることにする．

I 基本公式

$(z_1, \cdots, z_r)$ が r 次元の分布にしたがい，ある1組の非負の整数 $n_1, \cdots, n_r$ に対して，$0 \leqq m_1 \leqq n_1, \cdots, 0 \leqq m_r \leqq n_r$ なるすべての整数値 $m_1, \cdots, m_r$ について，$E[|z_1^{m_1} \cdots z_r^{m_r}|]$ が有限な値をもつとする．集合 $\{j_1, \cdots, j_p\}$ は1から r までのすべての整数値の集合 $\{1, \cdots, r\}$ の部分集合で，

$$\{k_1, \cdots, k_q\} = \{1, \cdots, r\} - \{j_1, \cdots, j_p\}$$

とする．ここで $p+q=r$ であり，また，$n = n_1 + \cdots + n_r$ とおく．このとき，$\phi(t_1, \cdots, t_r)$ を $(z_1, \cdots, z_r)$ の特性関数とすれば，

(1) $E[z_1^{n_1} \cdots z_r^{n_r} \operatorname{sgn}(z_{j_1} \cdots z_{j_p})]$

$$= \frac{1}{i^{p+n} \pi^p} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{dt_{j_1} \cdots dt_{j_p}}{t_{j_1} \cdots t_{j_p}} \left[\frac{\partial^n \phi(t_1, \cdots, t_r)}{\partial t_1^{n_1} \cdots \partial t_r^{n_r}} \right]_{t_{k_1} = \cdots = t_{k_q} = 0}$$

が成り立つ (Nabeya[5])．ここで $t_{j_1}, \cdots, t_{j_p}$ に関する積分は Cauchy の主値

$$\lim_{\substack{\varepsilon_1, \cdots, \varepsilon_p \to 0 \\ C_1, \cdots, C_p \to \infty}} \left(\int_{\varepsilon_1}^{C_1} + \int_{-C_1}^{-\varepsilon_1} \right) dt_{j_1} \cdots \left(\int_{\varepsilon_p}^{C_p} + \int_{-C_p}^{-\varepsilon_p} \right) dt_{j_p}$$

である．

(z_1, z_2, z_3, z_4) の分布は正規分布で，

$$E(z_i) = 0, \quad V(z_i) = 1, \quad \rho(z_i, z_j) = \rho_{ij} \quad (i \neq j)$$

とする．このとき (1) によって，つぎの諸公式が導かれる．

(2)　$E[|z_1|]=\sqrt{\dfrac{2}{\pi}}$

(3)　$E[|z_1z_2|]=\dfrac{2}{\pi}(\sqrt{1-\rho_{12}{}^2}+\rho_{12}\sin^{-1}\rho_{12})$

(4)　$E[z_1 \operatorname{sgn} z_2]=\sqrt{\dfrac{2}{\pi}}\rho_{12}$

(5)　$E[\operatorname{sgn}(z_1z_2)]=\dfrac{2}{\pi}\sin^{-1}\rho_{12}$

(6)　$E[z_1{}^2 \operatorname{sgn}(z_1z_2)]=\dfrac{2}{\pi}(\rho_{12}\sqrt{1-\rho_{12}{}^2}+\sin^{-1}\rho_{12})$

(7)　$E[|z_1|z_2 \operatorname{sgn}(z_3)]=\dfrac{2}{\pi}(\rho_{23}\sqrt{1-\rho_{13}{}^2}+\rho_{12}\sin^{-1}\rho_{13})$

(8)　$E[z_1z_2 \operatorname{sgn}(z_3z_4)]=\dfrac{2}{\pi}\Big\{\dfrac{1}{\sqrt{1-\rho_{34}{}^2}}(\rho_{13}\rho_{24}+\rho_{14}\rho_{23}$

$-\rho_{13}\rho_{23}\rho_{34}-\rho_{14}\rho_{24}\rho_{34})+\rho_{12}\sin^{-1}\rho_{34}\Big\}$

これらのうち，最後の3つは新しい結果であろう．(8) で $z_1=z_4$ とおけば (7) となり，また (8) で $z_1=z_2=z_3$ とおけば (6) が導かれる．

つぎに $E[\operatorname{sgn}(z_1z_2z_3z_4)]$ は一般の相関行列については数値積分を要する式になるが，ここでは後の便宜のため，その1つの形をあげておく．(z_1, z_2, z_3, z_4) の相関行列は

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 \end{pmatrix} \qquad (\text{ただし } \rho_{ij}=\rho_{ji})$$

であるが，これに対して，相関行列

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13}u & \rho_{14}u \\ \rho_{21} & 1 & \rho_{23}u & \rho_{24}u \\ \rho_{31}u & \rho_{32}u & 1 & \rho_{34} \\ \rho_{41}u & \rho_{42}u & \rho_{43} & 1 \end{pmatrix} \qquad (\text{ただし } 0\leqq u\leqq 1)$$

をもつ分布を考え，この分布で z_k と z_l を消去したときの z_i と z_j の偏相関係数を $\rho_{ij\cdot kl}(u)$ で表わす．このとき

$$(9)\quad E[\mathrm{sgn}\,(z_1z_2z_3z_4)]=\frac{4}{\pi^2}\Bigg\{\sin^{-1}\rho_{12}\sin^{-1}\rho_{34} +\int_0^1\frac{\rho_{13}}{\sqrt{1-\rho_{13}{}^2u^2}}\sin^{-1}\rho_{24\cdot13}(u)du +\int_0^1\frac{\rho_{14}}{\sqrt{1-\rho_{14}{}^2u^2}}\sin^{-1}\rho_{23\cdot14}(u)du +\int_0^1\frac{\rho_{23}}{\sqrt{1-\rho_{23}{}^2u^2}}\sin^{-1}\rho_{14\cdot23}(u)du +\int_0^1\frac{\rho_{24}}{\sqrt{1-\rho_{24}{}^2u^2}}\sin^{-1}\rho_{13\cdot24}(u)du\Bigg\}$$

となる（[5] 参照）．

ここでとくに2つの形の相関行列について，(9) を計算しておく．

相関行列が

$$(10)\quad \begin{pmatrix} 1 & \alpha & \beta & \alpha\beta \\ \alpha & 1 & \alpha\beta & \beta \\ \beta & \alpha\beta & 1 & \alpha \\ \alpha\beta & \beta & \alpha & 1 \end{pmatrix}$$

のときには，

$$\rho_{13\cdot24}(1)=\rho_{24\cdot13}(1)=\beta$$
$$\rho_{14\cdot23}(1)=\rho_{23\cdot14}(1)=-\alpha\beta$$

となる．ここで (10) の形からわかるように，$\rho_{ij\cdot kl}(1)$ の式で β のかわりに βu とおいたものが，$\rho_{ij\cdot kl}(u)$ となる．よって

$$\rho_{13\cdot24}(u)=\rho_{24\cdot13}(u)=\beta u$$
$$\rho_{14\cdot23}(u)=\rho_{23\cdot14}(u)=-\alpha\beta u$$

を得るから，(9) によって，

$$(11)\quad E[\mathrm{sgn}\,(z_1z_2z_3z_4)] =\frac{4}{\pi^2}\Bigg\{(\sin^{-1}\alpha)^2+2\int_0^1\frac{\beta}{\sqrt{1-\beta^2u^2}}\sin^{-1}(\beta u)du$$

$$-2\int_0^1 \frac{\alpha\beta}{\sqrt{1-\alpha^2\beta^2u^2}} \sin^{-1}(\alpha\beta u)du\Big\}$$

$$=\frac{4}{\pi^2}[(\sin^{-1}\alpha)^2+(\sin^{-1}\beta)^2-\{\sin^{-1}(\alpha\beta)\}^2]$$

となる.

つぎに相関行列が

$$(12)\quad \begin{pmatrix} 1 & \alpha & \beta & \alpha\beta \\ \alpha & 1 & \beta/\alpha & \beta \\ \beta & \beta/\alpha & 1 & \alpha \\ \alpha\beta & \beta & \alpha & 1 \end{pmatrix}$$

のときには,

$$\rho_{13\cdot24}(1)=\rho_{14\cdot23}(1)=\rho_{24\cdot13}(1)=0$$

$$\rho_{23\cdot14}(1)=\frac{\beta(1-\alpha^2)}{\alpha(1-\beta^2)}$$

となる. この場合にも, $\rho_{ij\cdot kl}(1)$ で β のかわりに βu とおけば, $\rho_{ij\cdot kl}(u)$ が得られる. よって,

$$\rho_{13\cdot24}(u)=\rho_{14\cdot23}(u)=\rho_{24\cdot13}(u)=0$$

$$\rho_{23\cdot14}(u)=\frac{\beta(1-\alpha^2)u}{\alpha(1-\beta^2u^2)}$$

となるから,

$$(13)\quad E[\mathrm{sgn}\,(z_1z_2z_3z_4)]$$

$$=\frac{4}{\pi^2}\Big\{(\sin^{-1}\alpha)^2+\int_0^1 \frac{\alpha\beta}{\sqrt{1-\alpha^2\beta^2u^2}} \sin^{-1}\frac{\beta(1-\alpha^2)u}{\alpha(1-\beta^2u^2)}du\Big\}$$

$$=\frac{4}{\pi^2}\Big\{(\sin^{-1}\alpha)^2+\int_0^{\alpha\beta} \frac{1}{\sqrt{1-u^2}} \sin^{-1}\frac{(1-\alpha^2)u}{\alpha^2-u^2}du\Big\}$$

が得られる.

II　正規分布の独立標本の場合

2変量正規分布の確率密度

$$(14)\quad f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\Big[-\frac{1}{2(1-\rho^2)}\Big\{\Big(\frac{x-\mu_1}{\sigma_1}\Big)^2$$

$$-2\rho\frac{x-\mu_1}{\sigma_1}\frac{y-\mu_2}{\sigma_2}+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\Big\}\Big]$$

をもつ互に独立な確率変数 $(x_1, y_1), \cdots, (x_n, y_n)$ の実現値に基づいて，相関係数 ρ を推定するものとする．

もしも (14) で ρ 以外の parameters $\mu_1, \mu_2, \sigma_1, \sigma_2$ がすべて既知ならば，$(x_i-\mu_1)/\sigma_1$ を改めて x_i とおき，$(y_i-\mu_2)/\sigma_2$ を改めて y_i とおくことにより，確率密度

$$\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right\}$$

が得られるから，ρ の推定のための尤度方程式

$$\text{(15)}\quad \frac{n\rho}{1-\rho^2}-\frac{\rho}{(1-\rho^2)^2}(\sum x_i^2-2\rho\sum x_iy_i+\sum y_i^2)+\frac{1}{1-\rho^2}\sum x_iy_i=0$$

が導かれる．ただし $\sum$ は本節では，特に指定のない限り，いずれも i について 1 から n までの和を示す．(15) は ρ について 3 次方程式であって，これを ρ について解いて得られる推定値を r_0 とすれば，$n\to\infty$ のとき，その分散は，

$$\text{(16)}\quad V(r_0)\sim\frac{1}{n}\frac{(1-\rho^2)^2}{1+\rho^2}$$

となることが知られている(Kendall and Stuart [4])．

しかし，この推定値は ρ 以外の parameters がすべて既知のときにしか利用できず，また (15) を解くことはめんどうなので，これによって ρ を推定したという数値例はあまりないであろう．

つぎに (14) のすべての parameters が未知のときに，これらの最尤推定値を求めると，ρ に対する値は

$$\text{(17)}\quad r_2=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2\}^{\frac{1}{2}}}$$

となって，これが最もふつうに使われる推定値である．r_2 については標本分布も正確に知れており，その分散は

$$\text{(18)}\quad V(r_2)=1-(1-\rho^2)\frac{n-2}{n-1}F\left(1, 1, \frac{n+1}{2}, \rho^2\right)$$

$$-\frac{\rho^2\Gamma^4\left(\frac{n}{2}\right)}{\Gamma^2\left(\frac{n-1}{2}\right)\Gamma^2\left(\frac{n+1}{2}\right)}F^2\left(\frac{1}{2},\frac{1}{2},\frac{n+1}{2},\rho^2\right)$$

$$\sim\frac{1}{n}(1-\rho^2)^2$$

となる (Kendall and Stuart [4]).　F は超幾何関数である.

ここで (17) の右辺の分子を $n-1$ で割れば，$\rho\sigma_1\sigma_2$ の不偏推定値となり，

$$\frac{1}{n-1}\sum(x_i-\bar{x})^2,\quad \frac{1}{n-1}\sum(y_i-\bar{y})^2$$

はそれぞれ $\sigma_1{}^2, \sigma_2{}^2$ の不偏推定値になっている．そこで σ_1, σ_2 が既知のときに，(17) の分母で $\sum(x_i-\bar{x})^2$, $\sum(y_i-\bar{y})^2$ をそれぞれ $(n-1)\sigma_1{}^2$, $(n-1)\sigma_2{}^2$ でおきかえることにより，ρ の推定値として，

$$(19)\quad r_1=\frac{1}{(n-1)\sigma_1\sigma_2}\sum(x_i-\bar{x})(y_i-\bar{y})$$

を用いることにすれば，r_2 より精度のよい推定値が得られるかも知れないと考えられる．しかしこれは間違いである．r_1 については

$$E(r_1)=\rho$$

$$(20)\quad V(r_1)=\frac{1}{n-1}(1+\rho^2)\sim\frac{1}{n}(1+\rho^2)$$

であって，少なくとも $n\to\infty$ としたときの漸近分散でみる限り，$\rho\neq0$ であれば

$$V(r_2)<V(r_1)$$

となる.

このことは Stuart [6] によって注意されているが，その理由は標本調査の理論でよく知られている比推定と不偏推定の精度の関係から説明することができる．いま，(A_n, B_n) の分布は，

$$E(A_n)=\rho,\quad E(B_n)\sim1$$

$$V(A_n)\sim\frac{1}{n}\sigma_{AA},\quad C(A_n,B_n)\sim\frac{1}{n}\sigma_{AB},\quad V(B_n)\sim\frac{1}{n}\sigma_{BB}$$

とすれば，A_n/B_n の分布については，

$$E\left(\frac{A_n}{B_n}\right)\sim\rho$$

$$(21)\quad V\left(\frac{A_n}{B_n}\right)\sim\frac{1}{n}(\sigma_{AA}-2\rho\sigma_{AB}+\rho^2\sigma_{BB})$$

が得られる．ここで，$\rho>0$ としよう．このとき，$2\rho\sigma_{AB}>\rho^2\sigma_{BB}$ が $V(A_n/B_n)<V(A_n)$ となるための必要十分条件であるが，これは，

$$(22)\quad \rho_{AB}=\frac{\sigma_{AB}}{\sigma_A\sigma_B}>\frac{\rho}{2}\frac{\sigma_B}{\sigma_A}$$

と同値である．ただし，$\sigma_A=\sqrt{\sigma_{AA}}$，$\sigma_B=\sqrt{\sigma_{BB}}$ とする．すなわち A_n と B_n の相関が，それらの標準偏差の比と，推定すべき値 ρ とから定まる一定値より大きければ，比推定のほうが不偏推定より精度がよい．

$$r_2=\frac{r_1}{s_1},\ \text{ただし}\ s_1=\frac{1}{(n-1)\sigma_1\sigma_2}\{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2\}^{\frac{1}{2}}$$

とおけば，

$$\sigma_A{}^2=\sigma_B{}^2=1+\rho^2,\ \rho_{AB}=\frac{2\rho}{1+\rho^2}$$

であって，(22) の不等式が成り立っている．$\rho<0$ のときも事情は同じである．

つぎに Huzii [2] が正規定常 Markov 過程の場合に提案した推定値 (III の (44)) に対応するつぎの推定値を考えてみる．

$$(23)\quad r_3=\sqrt{\frac{\pi}{2n(n-1)}}\frac{1}{\sigma_1}\sum(x_i-\bar{x})\mathrm{sgn}(y_i-\bar{y})$$

これは σ_1 が既知のときにしか利用できない．r_3 が ρ の不偏推定値であることを示し，その分散を計算しよう．

$$(24)\quad V(x_i-\bar{x})=\frac{n-1}{n}\sigma_1{}^2$$

$$\rho(x_i-\bar{x},y_i-\bar{y})=\rho$$

であるから，(4) により

$$E\{(x_i-\bar{x})\,\mathrm{sgn}\,(y_i-\bar{y})\}=\sqrt{\frac{n-1}{n}}\sigma_1\sqrt{\frac{2}{\pi}}\rho$$

となる．よって $E(r_3)=\rho$ は明らかである．

つぎに

$$\begin{aligned}V(r_3)&=\frac{\pi}{2n(n-1)}\frac{1}{\sigma_1{}^2}E[\{\sum(x_i-\bar{x})\mathrm{sgn}(y_i-\bar{y})\}^2]-\rho^2\\&=\frac{\pi}{2n(n-1)}\frac{1}{\sigma_1{}^2}E[\sum(x_i-\bar{x})^2\\&\qquad+\sum_{i\neq j}(x_i-\bar{x})(x_j-\bar{x})\,\mathrm{sgn}\,\{(y_i-\bar{y})(y_j-\bar{y})\}]-\rho^2\end{aligned}$$

を得る．ここで

$$E[\sum(x_i-\bar{x})^2]=(n-1)\sigma_1{}^2$$

である．また $i\neq j$ のとき，$x_i-\bar{x}$, $x_j-\bar{x}$ の分散はいずれも (24) の右辺で与えられ，$x_i-\bar{x}$, $x_j-\bar{x}$, $y_i-\bar{y}$, $y_j-\bar{y}$ の相関行列は

$$(25)\quad\begin{pmatrix}1 & -\frac{1}{n-1} & \rho & -\frac{\rho}{n-1}\\ -\frac{1}{n-1} & 1 & -\frac{\rho}{n-1} & \rho\\ \rho & -\frac{\rho}{n-1} & 1 & -\frac{1}{n-1}\\ -\frac{\rho}{n-1} & \rho & -\frac{1}{n-1} & 1\end{pmatrix}$$

となるから，(8) を用いて，

$$\begin{aligned}&E[(x_i-\bar{x})(x_j-\bar{x})\mathrm{sgn}\{(y_i-\bar{y})(y_j-\bar{y})\}]\\&=\frac{2}{\pi}\frac{n-1}{n}\sigma_1{}^2\left\{\rho^2\sqrt{1-\frac{1}{(n-1)^2}}+\frac{1}{n-1}\sin^{-1}\frac{1}{n-1}\right\}\\&=\frac{2}{\pi}\frac{\sigma_1{}^2}{n}\left\{\rho^2\sqrt{n(n-2)}+\sin^{-1}\frac{1}{n-1}\right\}\end{aligned}$$

となる．よって

$$(26)\quad\begin{aligned}V(r_3)&=\frac{\pi}{2n}+\frac{1}{n}\left\{\rho^2\sqrt{n(n-2)}+\sin^{-1}\frac{1}{n-1}\right\}-\rho^2\\&=\frac{\pi}{2n}-\left(1-\sqrt{1-\frac{2}{n}}\right)\rho^2+\frac{1}{n}\sin^{-1}\frac{1}{n-1}\sim\frac{1}{n}\left(\frac{\pi}{2}-\rho^2\right)\end{aligned}$$

となる．n が大のとき，$V(r_3)$ と $V(r_1)$ を比較すれば，

$$|\rho| \gtrless \sqrt{\frac{\pi}{4}-\frac{1}{2}}=0.53423$$

に応じて

$$V(r_1) \gtrless V(r_3)$$

となる．

つぎに r_3 を比推定値の形にした

$$(27)\quad r_4=\frac{\sum(x_i-\bar{x})\,\mathrm{sgn}\,(y_i-\bar{y})}{\sum|x_i-\bar{x}|}$$

について考えてみる．この推定値をつくるには，parameters はいっさい必要ない．ここで

$$r_4=\frac{r_3}{s_3},\quad \text{ただし}\ s_3=\sqrt{\frac{\pi}{2n(n-1)}}\frac{1}{\sigma_1}\sum|x_i-\bar{x}|$$

とおけば，$E(s_3)=1$ であって，$V(r_3)$ は (26) によって与えられ，$V(s_3)$ は (26) で $\rho=1$ とおくことによって与えられる．すなわち，

$$V(s_3)=\frac{\pi}{2n}-\left(1-\sqrt{1-\frac{2}{n}}\right)+\frac{1}{n}\sin^{-1}\frac{1}{n-1}\sim\frac{1}{n}\left(\frac{\pi}{2}-1\right)$$

となる．r_3 と s_3 の共分散は

$$\begin{aligned}C(r_3,\ s_3)&=\frac{\pi}{2n(n-1)}\frac{1}{\sigma_1{}^2}E\left[\sum|x_i-\bar{x}|\sum(x_i-\bar{x})\,\mathrm{sgn}\,(y_i-\bar{y})\right]-\rho\\&=\frac{\pi}{2n(n-1)}\frac{1}{\sigma_1{}^2}E\Big[\sum_i(x_i-\bar{x})^2\,\mathrm{sgn}\,\{(x_i-\bar{x})(y_i-\bar{y})\}\\&\qquad+\sum_{i\neq j}|x_i-\bar{x}|(x_j-\bar{x})\,\mathrm{sgn}\,(y_j-\bar{y})\Big]-\rho\end{aligned}$$

となる．ここで (6), (7) によって，

$$\begin{aligned}&E[(x_i-\bar{x})^2\,\mathrm{sgn}\,\{(x_i-\bar{x})(y_i-\bar{y})\}]\\&\quad=\frac{n-1}{n}\sigma_1{}^2\frac{2}{\pi}(\rho\sqrt{1-\rho^2}+\sin^{-1}\rho)\\&E[|x_i-\bar{x}|(x_j-\bar{x})\mathrm{sgn}(y_j-\bar{y})]\\&\quad=\frac{n-1}{n}\sigma_1{}^2\frac{2}{\pi}\left\{\rho\sqrt{1-\frac{\rho^2}{(n-1)^2}}+\frac{1}{n-1}\sin^{-1}\frac{\rho}{n-1}\right\}\end{aligned}$$

となるから，

$$C(r_3, s_3) = \frac{1}{n}(\rho\sqrt{1-\rho^2} + \sin^{-1}\rho)$$

$$+\frac{n-1}{n}\left\{\rho\sqrt{1-\frac{\rho^2}{(n-1)^2}} + \frac{1}{n-1}\sin^{-1}\frac{\rho}{n-1}\right\} - \rho$$

$$\sim\frac{1}{n}\{-\rho(1-\sqrt{1-\rho^2}) + \sin^{-1}\rho\}$$

が得られる．よって (21) から

$$(28)\qquad V(r_4) \sim \frac{1}{n}\left(\frac{\pi}{2} + \frac{\pi}{2}\rho^2 - 2\rho^2\sqrt{1-\rho^2} - 2\rho\sin^{-1}\rho\right)$$

となる．

つぎに

$$(29)\qquad r_5 = \frac{\sqrt{\pi}}{n(n-1)}\frac{1}{\sigma_1}\sum_{i<j}(x_i - x_j)\,\mathrm{sgn}\,(y_i - y_j)$$

について考える．これは σ_1 が既知のときにしか利用できない．r_1 または r_2 のように，$(x_1, \cdots, x_n)$ と $(y_1, \cdots, y_n)$ の 2 次形式と双 1 次形式からつくられる推定値については，

$$\sum_{i<j}(x_i - x_j)^2 = n\sum(x_i - \bar{x})^2$$

$$\sum_{i<j}(y_i - y_j)^2 = n\sum(y_i - \bar{y})^2$$

$$\sum_{i<j}(x_i - x_j)(y_i - y_j) = n\sum(x_i - \bar{x})(y_i - \bar{y})$$

によって，これらの左辺の和の形からつくられる推定値も，右辺の和の形からつくられる推定値も同じものとなる．さて，$i \neq j$ のとき，

$$V(x_i - x_j) = 2\sigma_1^2$$

$$\rho(x_i - x_j, y_i - y_j) = \rho$$

となるから，(4) によって

$$E[(x_i - x_j)\,\mathrm{sgn}\,(y_i - y_j)] = \sigma_1\frac{2}{\sqrt{\pi}}\rho$$

となり，

$$E(r_5) = \rho$$

は明らかである．また分散は

$$V(r_5)=\frac{\pi}{n^2(n-1)^2}\frac{1}{\sigma_1{}^2}E[\{\sum_{i<j}(x_i-x_j)\ \mathrm{sgn}\ (y_i-y_j)\}^2]-\rho^2$$
$$=\frac{\pi}{n^2(n-1)^2}\frac{1}{\sigma_1{}^2}\sum_{i<j}\sum_{k<l}E[(x_i-x_j)(x_k-x_l)$$
$$\mathrm{sgn}\ \{(y_i-y_j)(y_k-y_l)\}]-\rho^2$$

となるが，ここで右辺の各平均値を計算するのに，場合を分けて考えることにする．

1)　$i=k,\ j=l$ のときには，

$$E[(x_i-x_j)^2]=2\sigma_1{}^2$$

となり，この形の平均値は $\frac{1}{2}n(n-1)$ 個生ずる．

2)　$i=k,\ j\neq l$ のときには，$x_i-x_j,\ x_i-x_l,\ y_i-y_j,\ y_i-y_l$ の相関行列は

$$(30)\quad \begin{pmatrix} 1 & \frac{1}{2} & \rho & \frac{\rho}{2} \\ \frac{1}{2} & 1 & \frac{\rho}{2} & \rho \\ \rho & \frac{\rho}{2} & 1 & \frac{1}{2} \\ \frac{\rho}{2} & \rho & \frac{1}{2} & 1 \end{pmatrix}$$

となるから，(8) を用いて，

$$(31)\quad E[(x_i-x_j)(x_i-x_l)\ \mathrm{sgn}\ \{(y_i-y_j)(y_i-y_l)\}]$$
$$=\sigma_1{}^2\left(\frac{2\sqrt{3}}{\pi}\rho^2+\frac{1}{3}\right)$$

となる．

$i\neq k,\ j=l$ のときも (31) と同じ平均値を得る．

$i<j=k<l$ のときには，

$$E[(x_i-x_j)(x_j-x_l)\ \mathrm{sgn}\ \{(y_i-y_j)(y_j-y_l)\}]$$
$$=E[(x_j-x_i)(x_j-x_l)\ \mathrm{sgn}\ \{(y_j-y_i)(y_j-y_l)\}]$$

であって，$x_j-x_i,\ x_j-x_l,\ y_j-y_i,\ y_j-y_l$ の相関行列は (30) となるか

ら，この平均値も (31) に等しい．

$k<l=i<j$ のときも同様である．

これらの場合を合わせて，(31) の形の平均値が $n(n-1)(n-2)$ 個生ずる．

3) i, j, k, l のどれも異なるときには，

$$E[(x_i-x_j)(x_k-x_l)\,\mathrm{sgn}\,\{(y_i-y_j)(y_k-y_l)\}]$$
$$=E[(x_i-x_j)\,\mathrm{sgn}\,(y_i-y_j)]E[(x_k-x_l)\,\mathrm{sgn}\,(y_k-y_l)]$$
$$=\sigma_1{}^2\frac{4}{\pi}\rho^2$$

となり，この形の平均値が $\frac{1}{4}n(n-1)(n-2)(n-3)$ 個生ずる．

よって，

$$(32)\quad V(r_5)=\frac{4}{n^2(n-1)^2}\left\{\frac{n(n-1)}{4}\pi+n(n-1)(n-2)\left(\frac{\sqrt{3}}{2}\rho^2+\frac{\pi}{12}\right)\right.$$
$$\left.+\frac{1}{4}n(n-1)(n-2)(n-3)\rho^2\right\}-\rho^2$$
$$=\frac{1}{n(n-1)}\left\{\pi+(n-2)\left(2\sqrt{3}\,\rho^2+\frac{\pi}{3}\right)-(4n-6)\rho^2\right\}$$
$$\sim\frac{1}{n}\left\{\frac{\pi}{3}-(4-2\sqrt{3})\rho^2\right\}$$

となる．

つぎに r_5 を比推定値の形にした

$$(33)\quad r_6=\frac{\sum_{i<j}(x_i-x_j)\,\mathrm{sgn}\,(y_i-y_j)}{\sum_{i<j}|x_i-x_j|}$$

について考えてみる．この推定値をつくるには parameters はいっさい必要としない．ここで，

$$r_6=\frac{r_5}{s_5},\quad \text{ただし}\quad s_5=\frac{\sqrt{\pi}}{n(n-1)}\frac{1}{\sigma_1}\sum_{i<j}|x_i-x_j|$$

とおけば，$E(s_5)=1$ であって，$V(r_5)$ は (32) によって与えられ，$V(s_5)$ は (32) で $\rho=1$ とおいたものとなる．すなわち，

$$V(s_5)=\frac{1}{n(n-1)}\left\{\pi+(n-2)\left(2\sqrt{3}+\frac{\pi}{3}\right)-(4n-6)\right\}$$

$$\sim\frac{1}{n}\left(\frac{\pi}{3}-(4-2\sqrt{3})\right)$$

となる．r_5 と s_5 の共分散は

$$C(r_5, s_5)=\frac{\pi}{n^2(n-1)^2}\frac{1}{\sigma_1{}^2}E[\sum_{i<j}|x_i-x_j|\sum_{i<j}(x_i-x_j)\ \mathrm{sgn}\ (y_i-y_j)]$$
$$-\rho$$
$$=\frac{\pi}{n^2(n-1)^2}\frac{1}{\sigma_1{}^2}\sum_{i<j}\sum_{k<l}E[|x_i-x_j|(x_k-x_l)\ \mathrm{sgn}\ (y_k-y_l)]$$
$$-\rho$$

となる．ここでまた場合を分けて，右辺の各平均値を計算することにする．

1) $i=k,\ j=l$ のときには，(6) により，

$$E[(x_i-x_j)^2\ \mathrm{sgn}\ \{(x_i-x_j)(y_i-y_j)\}]$$
$$=2\sigma_1{}^2\frac{2}{\pi}(\rho\sqrt{1-\rho^2}+\sin^{-1}\rho)$$

となる．この形の平均値は $\dfrac{1}{2}n(n-1)$ 個生ずる．

2) $i=k,\ j\neq l$ のときには，$x_i-x_j,\ x_i-x_l,\ y_i-y_l$ の相関行列は

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{\rho}{2} \\ \frac{1}{2} & 1 & \rho \\ \frac{\rho}{2} & \rho & 1 \end{pmatrix}$$

となるから，(7) により

(34)　$$E[|x_i-x_j|(x_i-x_l)\ \mathrm{sgn}\ (y_i-y_l)]$$
$$=2\sigma_1{}^2\frac{2}{\pi}\left(\rho\sqrt{1-\frac{\rho^2}{4}}+\frac{1}{2}\sin^{-1}\frac{\rho}{2}\right)$$

となる．$i\neq k,\ j=l$ のとき，$i<j=k<l$ のとき，$k<l=i<j$ のときも，平均値は (34) と同じ値になる．この平均値が全部で $n(n-1)(n-2)$

個生ずる.

3)　i, j, k, l のどれも異なるときには,

$$E[|x_i-x_j|(x_k-x_l)\operatorname{sgn}(y_k-y_l)]$$
$$=E[|x_i-x_j|]E[(x_k-x_l)\operatorname{sgn}(y_k-y_l)]$$
$$=\sigma_1{}^2\frac{4}{\pi}\rho$$

となる．この形の平均値が $\dfrac{1}{4}n(n-1)(n-2)(n-3)$ 個生ずる.

よって

$$\begin{aligned}C(r_5, s_5)&=\frac{\pi}{n^2(n-1)^2}\left\{n(n-1)\frac{2}{\pi}(\rho\sqrt{1-\rho^2}+\sin^{-1}\rho)\right.\\&\quad+n(n-1)(n-2)\frac{4}{\pi}\left(\rho\sqrt{1-\frac{\rho^2}{4}}+\frac{1}{2}\sin^{-1}\frac{\rho}{2}\right)\\&\quad\left.+n(n-1)(n-2)(n-3)\frac{1}{\pi}\rho\right\}-\rho\\&=\frac{1}{n(n-1)}\left\{2(\rho\sqrt{1-\rho^2}+\sin^{-1}\rho)\right.\\&\quad+4(n-2)\left(\rho\sqrt{1-\frac{\rho^2}{4}}+\frac{1}{2}\sin^{-1}\frac{\rho}{2}\right)\\&\quad\left.-(4n-6)\rho\right\}\\&\sim\frac{1}{n}\left\{2\sin^{-1}\frac{\rho}{2}-4\rho\left(1-\sqrt{1-\frac{\rho^2}{4}}\right)\right\}\end{aligned}$$

が得られるから，(21) により,

$$(35)\quad V(r_6)\sim\frac{1}{n}\left(\frac{\pi}{3}+\frac{\pi}{3}\rho^2+4\sqrt{3}\,\rho^2-4\rho^2\sqrt{4-\rho^2}-4\rho\sin^{-1}\frac{\rho}{2}\right)$$

となる.

つぎに

$$(36)\quad r_7=\sin\left[\frac{\pi}{2n}\sum\operatorname{sgn}\{(x_i-\bar{x})(y_i-\bar{y})\}\right]$$

について考える．この後で Kendall の順位相関係数にもとづく推定値 r_8 を取り扱うが，推定値の形の上では，r_7 と r_8 の関係が，r_3 と

r_5 の関係，ないしは r_4 と r_6 の関係と同じになっている．

$\sin^{-1} r_7$ については，(5) により

$$E[\sin^{-1} r_7]=\sin^{-1}\rho$$

となることは明らかである．また分散は

$$V(\sin^{-1} r_7)=\frac{\pi^2}{4n^2}E[(\sum \mathrm{sgn}\,\{(x_i-\bar{x})(y_i-\bar{y})\})^2]-(\sin^{-1}\rho)^2$$
$$=\frac{\pi^2}{4n^2}(n+\sum_{i\neq j}E[\mathrm{sgn}\,\{(x_i-\bar{x})(x_j-\bar{x})(y_i-\bar{y})(y_j-\bar{y})\}])-(\sin^{-1}\rho)^2$$

となる．ここで $x_i-\bar{x}$, $x_j-\bar{x}$, $y_i-\bar{y}$, $y_j-\bar{y}$ の相関行列は (25) であったが，この行列は $\alpha=-\dfrac{1}{n-1}$, $\beta=\rho$ とおいたとき，(10) の形になっている．よって (11) から

$$E[\mathrm{sgn}\,\{(x_i-\bar{x})(x_j-\bar{x})(y_i-\bar{y})(y_j-\bar{y})\}]$$
$$=\frac{4}{\pi^2}\left\{(\sin^{-1}\rho)^2+\left(\sin^{-1}\frac{1}{n-1}\right)^2-\left(\sin^{-1}\frac{\rho}{n-1}\right)^2\right\}$$

を得る．したがって

$$V(\sin^{-1} r_7)=\frac{\pi^2}{4n^2}\left[n+n(n-1)\frac{4}{\pi^2}\left\{(\sin^{-1}\rho)^2+\left(\sin^{-1}\frac{1}{n-1}\right)^2-\left(\sin^{-1}\frac{\rho}{n-1}\right)^2\right\}\right]-(\sin^{-1}\rho)^2$$
$$=\frac{1}{n}\left\{\frac{\pi^2}{4}-(\sin^{-1}\rho)^2\right\}+\frac{n-1}{n}\left\{\left(\sin^{-1}\frac{1}{n-1}\right)^2-\left(\sin^{-1}\frac{\rho}{n-1}\right)^2\right\}$$
$$\sim\frac{1}{n}\left\{\frac{\pi^2}{4}-(\sin^{-1}\rho)^2\right\}$$

となる．よって r_7 については，

$$V(r_7)\sim\frac{1}{n}(1-\rho^2)\left\{\frac{\pi^2}{4}-(\sin^{-1}\rho)^2\right\} \tag{37}$$

となる．

最後に

(38)　$r_8 = \sin\left[\frac{\pi}{n(n-1)}\sum_{i<j} \mathrm{sgn}\,\{(x_i-x_j)(y_i-y_j)\}\right]$

とすれば，$\frac{2}{\pi}\sin^{-1} r_8$ が Kendall の順位相関係数となる．これについては，

$$E\left(\frac{2}{\pi}\sin^{-1} r_8\right)=\frac{2}{\pi}\sin^{-1}\rho$$

$$V\left(\frac{2}{\pi}\sin^{-1} r_8\right)=\frac{2}{n(n-1)}\left[1-\left(\frac{2}{\pi}\sin^{-1}\rho\right)^2 +2(n-2)\left\{\frac{1}{9}-\left(\frac{2}{\pi}\sin^{-1}\frac{\rho}{2}\right)^2\right\}\right]$$

$$\sim\frac{1}{n}\left\{\frac{4}{9}-\left(\frac{4}{\pi}\sin^{-1}\frac{\rho}{2}\right)^2\right\}$$

となることが知られている（Kendall [3]）．よってこれから

(39)　$V(r_8)\sim\frac{1}{n}(1-\rho^2)\left\{\frac{\pi^2}{9}-4\left(\sin^{-1}\frac{\rho}{2}\right)^2\right\}$

を得る．

以上によって，2変量正規分布の場合について，相関係数 ρ の 9 つの推定値 r_i の分散が，$n\to\infty$ のときにいずれも $\sim c_i(\rho)/n$ の形となることがわかった．これらの $c_i(\rho)$ の値を $\rho=0(0.02)1$ に対して図示したものが 24 頁の図である．r_i に対する $c_i(\rho)$ の値をこの図では数字 i で示してある．この図からつぎの事実が読みとれる．

1)　$c_1(\rho)$ と $c_2(\rho)$ の組，$c_3(\rho)$ と $c_4(\rho)$ の組，$c_5(\rho)$ と $c_6(\rho)$ の組は，いずれも対応する不偏推定値と比推定値に関する値であるが，この各組の値は $\rho=0$ のときには一致して，ρ が大きくなるにしたがって離れている．どの組についても比推定のほうが精度がよく，比推定については $\rho\to1$ のとき $c_i(\rho)\to0$ となるが，不偏推定の場合には $c_i(\rho)\to0$ とはならない．とくに $c_1(\rho)$ は ρ が大きくなるにしたがって増大していく．

2)　$c_2(\rho)$, $c_6(\rho)$, $c_8(\rho)$ の線はほとんど同じ傾向で進んでおり，$c_0(\rho)$ の線に比較的よく（差でみる限り）接近している．

3）r_7 はその形からみて，精度のよくないことは容易に想像がつくが，ρ が 1 に近くなれば，不偏推定値のグループより精度はよくなる．

4）Huzii [2] は正規定常 Markov 過程の場合について，r_1 の形のものと r_3 の形のものとの間で精度を比較しているが，分散が既知であっても，不偏推定を用いるより比推定を用いるほうが精度がよいので，この 2 つの比較だけでは片手落ちである．そして，r_2 と r_4 の比較では，やはり r_2 のほうがすぐれている．

III　正規定常 Markov 過程の場合

つぎに

$$x_t-\rho x_{t-1}=u_t \quad (t=0, \pm 1, \pm 2, \cdots)$$

で与えられる正規定常 Markov 過程 $\{x_t\}$ の場合を取り扱うことにする．ただし $|\rho|<1$ であって，$\{u_t\}$ は平均値が 0 の正規独立定常過程とする．ここで，

$$V(x_t)=\frac{V(u_t)}{1-\rho^2}=\sigma^2$$

とおく．この場合には，x_t と x_{t-h} の相関係数を ρ_h とすれば，$\rho_h=\rho^h$ となる．しかしこの確率過程の型が未知のときには，ρ_1 の推定値を算出して，これを h 乗して ρ_h の推定値を求めるわけにはいかない．ここでは $n+h$ 個の引き続いた観測値 $x_1, \cdots, x_{n+h}$ が与えられたとき，これを平均値が 0 の定常過程の実現値とみなして，前節にならって ρ_h のいろいろな推定値をつくる．そしてそれらの分散を正規定常 Markov 過程の場合に計算してみることにする．

まず σ が既知として

$$r_1=\frac{1}{n\sigma^2}\sum_{i=1}^{n} x_i x_{i+h} \tag{40}$$

を考えると，

$$E(r_1)=\rho^h$$

$$(41)\quad V(r_1)=\frac{1}{n^2}\left\{n(1+\rho^{2h})+4\sum_{k=1}^{n-1}(n-k)\rho^{2k}\right.$$
$$\left.+2\sum_{k=1}^{h-1}(n-k)(\rho^{2h}-\rho^{2k})\right\}$$
$$\sim\frac{1}{n}\left\{\frac{(1+\rho^2)(1+\rho^{2h})}{1-\rho^2}+2h\rho^{2h}\right\}$$

となることは容易にわかる (Bartlett [1] 参照).

もっともよく使われる推定値は, r_1 に対応する比推定値

$$(42)\quad r_2=\frac{n+h}{n}\frac{\sum_{i=1}^{n}x_ix_{i+h}}{\sum_{i=1}^{n+h}x_i^2}$$

であって, この値は σ の値を知らなくても計算できる. これについても Bartlett [1] によって,

$$(43)\quad V(r_2)\sim\frac{1}{n}\left\{\frac{(1+\rho^2)(1-\rho^{2h})}{1-\rho^2}-2h\rho^{2h}\right\}$$

が得られている.

Huzii [2] による

$$(44)\quad r_3=\sqrt{\frac{\pi}{2}}\frac{1}{n\sigma}\sum_{i=1}^{n}x_i\,\mathrm{sgn}\,(x_{i+h})$$

については, (4) によって

$$E(r_3)=\rho^h$$

であり, また分散は

$$V(r_3)=\frac{\pi}{2}\frac{1}{n^2\sigma^2}E\left[\left\{\sum_{i=1}^{n}x_i\,\mathrm{sgn}\,(x_{i+h})\right\}^2\right]-\rho^{2h}$$
$$=\frac{\pi}{2}\frac{1}{n^2\sigma^2}\left\{\sum_{i=1}^{n}E[x_i^2]+\sum_{\substack{i,j=1\\ i\neq j}}^{n}E[x_ix_j\,\mathrm{sgn}\,(x_{i+h}x_{j+h})]\right\}-\rho^{2h}$$
$$=\frac{\pi}{2}\frac{1}{n^2\sigma^2}\left\{n\sigma^2+2\sum_{k=1}^{n-1}(n-k)E[x_ix_{i+k}\,\mathrm{sgn}\,(x_{i+h}x_{i+h+k})]\right\}$$
$$-\rho^{2h}$$

となる. ここで $(x_i, x_{i+k}, x_{i+h}, x_{i+h+k})$ の相関行列は

$$(45)\quad \begin{pmatrix} 1 & \rho^k & \rho^h & \rho^{h+k} \\ \rho^k & 1 & \rho^{|k-h|} & \rho^h \\ \rho^h & \rho^{|k-h|} & 1 & \rho^k \\ \rho^{h+k} & \rho^h & \rho^k & 1 \end{pmatrix}$$

であるから，(8) によって，

$$E[x_i x_{i+k}\,\mathrm{sgn}\,(x_{i+h}x_{i+h+k})]$$
$$=\sigma^2\frac{2}{\pi}(\rho^{2h}\sqrt{1-\rho^{2k}}+\rho^k\sin^{-1}\rho^k)$$

となる．よって

$$(46)\quad V(r_3)=\frac{1}{n^2}\left\{\frac{\pi}{2}n+2\sum_{k=1}^{n-1}(n-k)(\rho^{2h}\sqrt{1-\rho^{2k}}+\rho^k\sin^{-1}\rho^k)\right\}$$
$$-\rho^{2h}$$
$$=\frac{1}{n^2}\left\{n\left(\frac{\pi}{2}-\rho^{2h}\right)-2\rho^{2h}\sum_{k=1}^{n-1}(n-k)(1-\sqrt{1-\rho^{2k}})\right.$$
$$\left.+2\sum_{k=1}^{n-1}(n-k)\rho^k\sin^{-1}\rho^k\right\}$$
$$\sim\frac{1}{n}\left\{\frac{\pi}{2}-\rho^{2h}-2\rho^{2h}\sum_{k=1}^{\infty}(1-\sqrt{1-\rho^{2k}})+2\sum_{k=1}^{\infty}\rho^k\sin^{-1}\rho^k\right\}$$

を得る．(46) の第 3 辺は Huzii [2] によって，ρ に関する級数展開の形で与えられている．

つぎに r_3 を比推定の形にした

$$(47)\quad r_4=\frac{n+h}{n}\frac{\sum_{i=1}^{n}x_i\,\mathrm{sgn}\,(x_{i+h})}{\sum_{i=1}^{n+h}|x_i|}$$

について考えてみる．ここで

$$r_4=\frac{r_3}{s_3},\quad \text{ただし}\ s_3=\sqrt{\frac{\pi}{2}}\frac{1}{(n+h)\sigma}\sum_{i=1}^{n+h}|x_i|$$

とおけば，$E(s_3)=1$ であって，$V(r_3)$ は (46) によって与えられ，$V(s_3)$ は (46) で h を 0 とおいた後，n を $n+h$ でおきかえたものとなる．すなわち，

$$V(s_3)=\frac{1}{(n+h)^2}\left\{(n+h)\left(\frac{\pi}{2}-1\right)\right.$$
$$-2\sum_{k=1}^{n+h-1}(n+h-k)(1-\sqrt{1-\rho^{2k}})$$
$$\left.+2\sum_{k=1}^{n+h-1}(n+h-k)\rho^k\sin^{-1}\rho^k\right\}$$
$$\sim\frac{1}{n}\left\{\frac{\pi}{2}-1-2\sum_{k=1}^{\infty}(1-\sqrt{1-\rho^{2k}})+2\sum_{k=1}^{\infty}\rho^k\sin^{-1}\rho^k\right\}$$

である．また (7) によって

$$C(r_3,s_3)=\frac{\pi}{2}\frac{1}{n(n+h)\sigma^2}E\left[\sum_{i=1}^{n+h}|x_i|\sum_{i=1}^{n}x_i\,\mathrm{sgn}\,(x_{i+h})\right]-\rho^h$$
$$=\frac{\pi}{2}\frac{1}{n(n+h)\sigma^2}\left\{\sum_{k=1}^{n-1}(n-k)E[|x_{i-k}|x_i\,\mathrm{sgn}\,(x_{i+h})]\right.$$
$$+\sum_{k=0}^{h}nE[|x_{i+k}|x_i\,\mathrm{sgn}\,(x_{i+h})]$$
$$\left.+\sum_{k=1}^{n-1}(n-k)E[|x_{i+h+k}|x_i\,\mathrm{sgn}\,(x_{i+h})]\right\}-\rho^h$$
$$=\frac{1}{n(n+h)}\left\{\sum_{k=1}^{n-1}(n-k)(\rho^h\sqrt{1-\rho^{2h+2k}}+\rho^k\sin^{-1}\rho^{h+k})\right.$$
$$+\sum_{k=0}^{h}n(\rho^h\sqrt{1-\rho^{2h-2k}}+\rho^k\sin^{-1}\rho^{h-k})$$
$$\left.+\sum_{k=1}^{n-1}(n-k)(\rho^h\sqrt{1-\rho^{2k}}+\rho^{h+k}\sin^{-1}\rho^k)\right\}-\rho^h$$
$$\sim\frac{1}{n}\left\{\frac{\pi}{2}\rho^h-\rho^h-2\rho^h\sum_{k=1}^{\infty}(1-\sqrt{1-\rho^{2k}})\right.$$
$$+\sum_{k=1}^{\infty}\rho^k\sin^{-1}\rho^{h+k}+\sum_{k=1}^{\infty}\rho^{h+k}\sin^{-1}\rho^k$$
$$\left.+\sum_{k=0}^{h-1}\rho^k\sin^{-1}\rho^{h-k}\right\}$$

を得る．よって (21) から

$$(48)\qquad V(r_4)\sim\frac{1}{n}\left\{\frac{\pi}{2}(1-\rho^{2h})+2\sum_{k=1}^{h}(\rho^k-\rho^{2h-k})\sin^{-1}\rho^k\right\}$$

を得る．

最後に

$$(49)\quad r_7=\sin\left\{\frac{\pi}{2n}\sum_{i=1}^{n}\operatorname{sgn}(x_i x_{i+h})\right\}$$

とすれば，(5) によって

$$E(\sin^{-1} r_7)=\sin^{-1}\rho^h$$

$$V(\sin^{-1} r_7)=\frac{\pi^2}{4n^2}E\left[\left\{\sum_{i=1}^{n}\operatorname{sgn}(x_i x_{i+h})\right\}^2\right]-(\sin^{-1}\rho^h)^2$$

$$=\frac{\pi^2}{4n^2}\left\{n+2\sum_{k=1}^{n-1}(n-k)E[\operatorname{sgn}(x_i x_{i+h} x_{i+k} x_{i+h+k})]\right\}-(\sin^{-1}\rho^h)^2$$

となる．ここで $(x_i, x_{i+k}, x_{i+h}, x_{i+h+k})$ の相関行列は (45) であったが，この相関行列は $h>k$ のときには (12) で $\alpha=\rho^k, \beta=\rho^h$ とおいたもの，$h<k$ のときには $(x_i, x_{i+h}, x_{i+k}, x_{i+h+k})$ の相関行列が (12) で $\alpha=\rho^h$，$\beta=\rho^k$ とおいたものになる．よって $i<j$ のとき

$$f(i,j,\rho)=\int_0^{\rho^{i+j}}\frac{1}{\sqrt{1-u^2}}\sin^{-1}\frac{(1-\rho^{2i})u}{\rho^{2i}-u^2}du$$

$$\sim\frac{\rho^{2j}(1-\rho^{2i})}{2}\quad(j\to\infty)$$

とおけば，(13) から，$h>k$ のときには

$$E[\operatorname{sgn}(x_i x_{i+h} x_{i+k} x_{i+h+k})]$$

$$=\frac{4}{\pi^2}\{(\sin^{-1}\rho^k)^2+f(k,h,\rho)\}$$

$h<k$ のときには

$$E[\operatorname{sgn}(x_i x_{i+h} x_{i+k} x_{i+h+k})]$$

$$=\frac{4}{\pi^2}\{(\sin^{-1}\rho^h)^2+f(h,k,\rho)\}$$

となる．また $h=k$ のときには，(5) によってこの平均値は

$$E[\operatorname{sgn}(x_i x_{i+2h})]=\frac{2}{\pi}\sin^{-1}\rho^{2h}$$

となる．よって

$$V(\sin^{-1} r_7)=\frac{1}{n^2}\Bigg[\frac{\pi^2}{4}n+\pi(n-h)\sin^{-1}\rho^{2h}$$
$$+2\sum_{k=1}^{h-1}(n-k)\{(\sin^{-1}\rho^k)^2+f(k,h,\rho)\}$$
$$+2\sum_{k=h+1}^{n-1}(n-k)\{(\sin^{-1}\rho^h)^2+f(h,k,\rho)\}\Bigg]$$
$$-(\sin^{-1}\rho^h)^2$$
$$\sim\frac{1}{n}\Bigg\{\frac{\pi^2}{4}+\pi\sin^{-1}\rho^{2h}-(2h+1)(\sin^{-1}\rho^h)^2$$
$$+2\sum_{k=1}^{h-1}(\sin^{-1}\rho^k)^2+2\sum_{k=1}^{h-1}f(k,h,\rho)$$
$$+2\sum_{k=h+1}^{\infty}f(h,k,\rho)\Bigg\}$$

を得る．したがって

$$(50)\quad V(r_7)\sim\frac{1}{n}(1-\rho^{2h})\Bigg\{\frac{\pi^2}{4}+\pi\sin^{-1}\rho^{2h}-(2h+1)(\sin^{-1}\rho^h)^2$$
$$+2\sum_{k=1}^{h-1}(\sin^{-1}\rho^k)^2+2\sum_{k=1}^{h-1}f(k,h,\rho)+2\sum_{k=h+1}^{\infty}f(h,k,\rho)\Bigg\}$$

となる．

とくに (41), (43), (46), (48), (50) の漸近式は，$h=1$ のときにはつぎの形となる．

$$V(r_1)\sim\frac{1}{n}\left\{\frac{(1+\rho^2)^2}{1-\rho^2}+2\rho^2\right\}$$

$$V(r_2)\sim\frac{1}{n}(1-\rho^2)$$

$$V(r_3)\sim\frac{1}{n}\left\{\frac{\pi}{2}-\rho^2+2\sum_{k=1}^{\infty}\rho^k\sin^{-1}\rho^k-2\rho^2\sum_{k=1}^{\infty}(1-\sqrt{1-\rho^{2k}})\right\}$$

$$V(r_4)\sim\frac{1}{n}\frac{\pi}{2}(1-\rho^2)$$

$$V(r_7)\sim\frac{1}{n}(1-\rho^2)\Bigg\{\frac{\pi^2}{4}+\pi\sin^{-1}\rho^2-3(\sin^{-1}\rho)^2$$
$$+2\sum_{k=2}^{\infty}\int_0^{\rho^{k+1}}\frac{1}{\sqrt{1-u^2}}\sin^{-1}\frac{(1-\rho^2)u}{\rho^2-u^2}du\Bigg\}$$

以上で考えた r_1, r_2, r_3, r_4, r_7 はいずれも前節で同じ記号で表わした推定値に対応するものである．これらの分散を図に示したのが 25 頁の図であって，不偏推定値と対応する比推定値の分散の大小，$V(r_1)$ と $V(r_3)$ の大小などについての関係は，前節の場合とほぼ同様である．24 頁の図では ρ が大きくなるにしたがって，$c_1(\rho)$ は単調増加，$c_3(\rho)$ は単調減少であって，$\rho\to 1$ のときに両者はいずれも極限値をもっていたが，25 頁の図では $c_1(\rho)$, $c_3(\rho)$ はともに単調増加で，$\rho\to 1$ のとき ∞ に発散する．なお本節の場合には，r_5, r_6, r_8 に対応する推定値をつくることはできない．

参考文献

[1] Bartlett, M. S., On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, Vol. 8, No. 1, 1946.

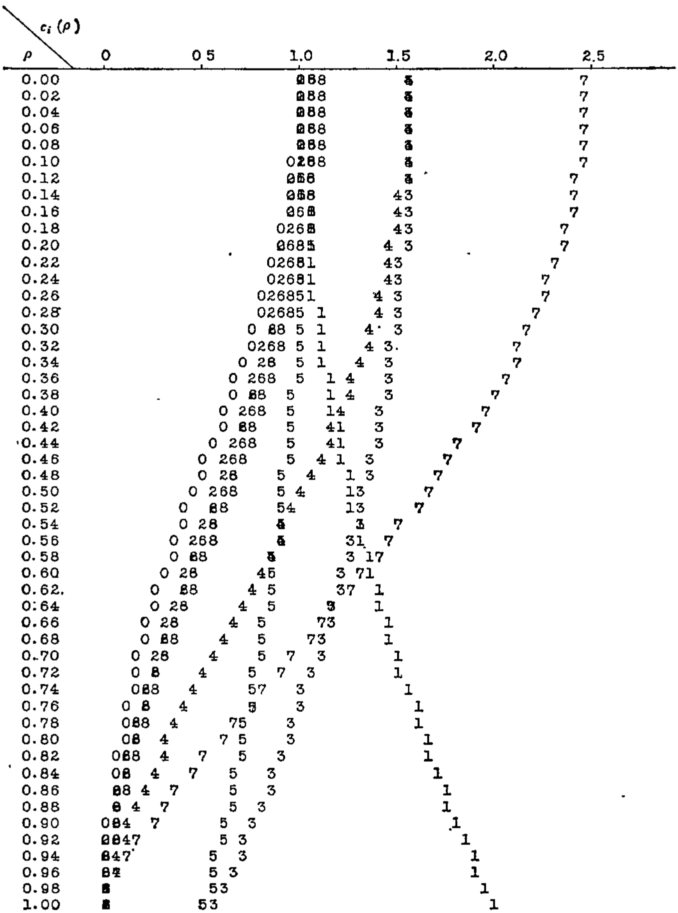
[2] Huzii, M., On a simplified method of the estimation of the correlogram for a stationary Gaussian process. *Annals of the Institute of Statistical Mathematics*, Vol. 14, No. 3, 1963.

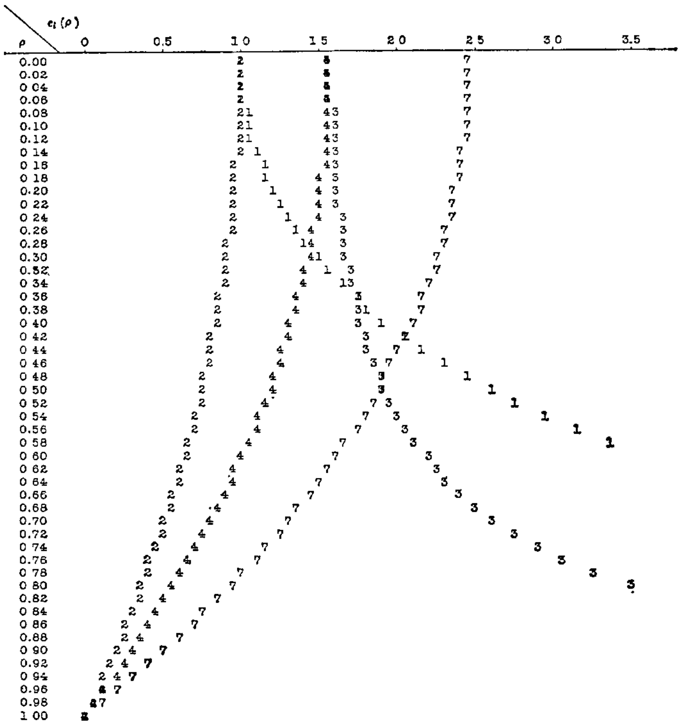
[3] Kendall, M. G., *Rank Correlation Methods*. Griffin, 1955.

[4] Kendall, M. G. and Stuart A., *The Advanced Theory of Statistics*, Vol. 1 and 2. Griffin, 1958 and 1961.

[5] Nabeya, S., Absolute and incomplete moments of the multivariate normal distribution. *Biometrika*, Vol. 48, No. 1 and 2, 1961.

[6] Stuart, A., A paradox in statistical estimation. *Biometrika*, Vol. 42, No. 1 and 2, 1955.


$c_i(\rho)$
ρ
0
0.5
1.0
1.5
2.0
2.5


$e_i(\rho)$
ρ
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5