2次元正規分布モデルにおける相関係数の多重比較法

2019SS088 安田竜規　2019SS095 横山颯

指導教員：白石高章

1　はじめに

本論文では, k 標本の 2 次元正規分布モデルにおける相関係数の多重比較検定法を考察する。また得られた検定方式から, 実際のデータに対する解析を行う.

2　モデルの設定

k 群モデルで第 i 群の 2 次元標本 $\begin{pmatrix} X_{i1} \\ Y_{i1} \end{pmatrix}, \cdots, \begin{pmatrix} X_{in_i} \\ Y_{in_i} \end{pmatrix}$ は, 平均が $\boldsymbol{\mu}_i = \begin{pmatrix} \mu_{1i} \\ \mu_{2i} \end{pmatrix}$, 分散共分散行列が $\boldsymbol{\Sigma}_i = \begin{pmatrix} \sigma_{1i}^2 & \rho_i\sigma_{1i}\sigma_{2i} \\ \rho_i\sigma_{1i}\sigma_{2i} & \sigma_{2i}^2 \end{pmatrix}$ である 2次元正規分布に従い, $j = 1, \cdots, n_i,\ i = 1, \cdots, k$ において $\begin{pmatrix} X_{ij} \\ Y_{ij} \end{pmatrix} \sim \mathrm{N}_2(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ は互いに独立であるものとする.

表 1　k 群モデル

群	サイズ	データ
第 1 群	n_1	$\begin{pmatrix} X_{11} \\ Y_{11} \end{pmatrix}, \cdots, \begin{pmatrix} X_{1n_1} \\ Y_{1n_1} \end{pmatrix}$
第 2 群	n_2	$\begin{pmatrix} X_{21} \\ Y_{21} \end{pmatrix}, \cdots, \begin{pmatrix} X_{2n_2} \\ Y_{2n_2} \end{pmatrix}$
$\vdots$	$\vdots$	$\vdots$
第 k 群	n_k	$\begin{pmatrix} X_{k1} \\ Y_{k1} \end{pmatrix}, \cdots, \begin{pmatrix} X_{kn_k} \\ Y_{kn_k} \end{pmatrix}$

比較のための検定は, $i = 1, \cdots, k$ に対して, ある $\rho_{0i} \in (-1, 1)$ を与えたときに,

$$\text{帰無仮説 } H_{0i}\text{: } \rho_i = \rho_{0i}$$

に対して 3 種の対立仮説

① 両側対立仮説 $H_{0i}^{A\pm}$: $\rho_i \neq \rho_{0i}$

② 片側対立仮説 H_{0i}^{A+} : $\rho_i > \rho_{0i}$

③ 片側対立仮説 H_{0i}^{A-} : $\rho_i < \rho_{0i}$

となる. サイズ全体の集合を $N \equiv \{n_i | 1 \leq i \leq k\}$, 総標本サイズを $n' \equiv \sum_{i=1}^{k} n_i$ とする.

3　統計量の漸近分布

統計量の漸近分布を求める.
確率変数列 $\{Y_n\}$ が定数 c に確率収束することと, 確率変数列 $\{Z_n\}$ が確率変数 Z に分布収束することをそれぞれ, 記号 $Y_n \xrightarrow{P} c$, $Z_n \xrightarrow{\mathcal{L}} Z$ で表す.

補題 1 n 個の確率ベクトル $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \cdots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ について, これらは互いに独立であり, 平均が

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},$$

分散共分散行列が

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

であるような 2 次元正規分布に従うとする.
このとき, 標本相関係数を

$$\hat{\rho} \equiv \frac{1/n \cdot \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{1/n \cdot \sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{1/n \cdot \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

と置いたときに, $n \to \infty$ として

$$\sqrt{n}(\hat{\rho} - \rho) \xrightarrow{\mathcal{L}} \mathrm{N}\left(0, (1 - \rho^2)^2\right)$$

が成り立つ.

証明　標本標準偏差 $\hat{\sigma}_X$, $\hat{\sigma}_Y$, 標本共分散 $\hat{\sigma}_{XY}$ を

$$\hat{\sigma}_X \equiv \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2},\quad \hat{\sigma}_Y \equiv \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2},$$

$$\hat{\sigma}_{XY} \equiv \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

とする. $\sqrt{n}(\hat{\rho} - \rho)$ は, $\hat{\sigma}_X$, $\hat{\sigma}_Y, \hat{\sigma}_{XY}$ を使って表すと,

$$\sqrt{n}(\hat{\rho} - \rho) = \frac{1}{\hat{\sigma}_X\hat{\sigma}_Y}\sqrt{n}\left(\hat{\sigma}_{XY} - \rho\right) + \frac{\rho}{\hat{\sigma}_X}\sqrt{n}\left(\frac{1}{\hat{\sigma}_Y} - 1\right) + \rho\sqrt{n}\left(\frac{1}{\hat{\sigma}_X} - 1\right) \tag{1}$$

となる. ここで, X_i, Y_i を標準化したもの

$$X_i' \equiv \frac{X_i - \mu_1}{\sigma_1},\ Y_i' \equiv \frac{Y_i - \mu_2}{\sigma_2}$$

について考える.

$$\begin{pmatrix} X_i' \\ Y_i' \end{pmatrix} \sim \mathrm{N}_2\left(\mathbf{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

であり,

$$\hat{\rho} = \frac{1/n \cdot \sum_{i=1}^{n}(X_i' - \bar{X}')(Y_i' - \bar{Y}')}{\sqrt{1/n \cdot \sum_{i=1}^{n}(X_i' - \bar{X}')^2}\sqrt{1/n \cdot \sum_{i=1}^{n}(Y_i' - \bar{Y}')^2}}$$

と表せるので, 一般性を失うことなく

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} = \begin{pmatrix} X_i' \\ Y_i' \end{pmatrix} \sim \mathrm{N}_2\left(\mathbf{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

$(i = 1, \cdots, n)$ としてよい.

ここで, $\mathrm{N}_2\left(\mathbf{0}, \boldsymbol{I}_2\right)$ に従う n 個の確率ベクトル $\begin{pmatrix} Z_{11} \\ Z_{21} \end{pmatrix}, \cdots, \begin{pmatrix} Z_{1n} \\ Z_{2n} \end{pmatrix}$ を与えたときに, X_i と Y_i は, $i = 1, \cdots, k$ に対して

$$X_i = Z_{1i},\ Y_i = \rho Z_{1i} + \sqrt{1-\rho^2} Z_{2i} \tag{2}$$

と表現できる.

ここで, 関数 $g(x) = \dfrac{1}{\sqrt{x}}$ を与えると, $g'(x) = -\dfrac{1}{2x^{3/2}}$ であり, 平均値の定理から,

$$|a_n - 1| < |\hat{\sigma}_X^2 - 1|, \quad |b_n - 1| < |\hat{\sigma}_Y^2 - 1|$$

を満たす $a_n,\ b_n$ が

$$g'(a_n) = \frac{g(\hat{\sigma}_X^2) - g(1)}{\hat{\sigma}_X^2 - 1} \Rightarrow g'(a_n) = \frac{1/\hat{\sigma}_X - 1}{\hat{\sigma}_X^2 - 1}\ , \tag{3}$$

$$g'(b_n) = \frac{g(\hat{\sigma}_Y^2) - g(1)}{\hat{\sigma}_Y^2 - 1} \Rightarrow g'(b_n) = \frac{1/\hat{\sigma}_Y - 1}{\hat{\sigma}_Y^2 - 1} \tag{4}$$

となるように存在する.

次に, $1/\hat{\sigma}_X$ と $1/(\hat{\sigma}_X\hat{\sigma}_Y)$ の確率収束先を調べる. 大数の法則から $\hat{\sigma}_X^2 \xrightarrow{P} \mathrm{V}(X_i) = 1$ が成り立つことと, $g(x)$ は連続関数であることから,

$$g(\hat{\sigma}_X^2) = \frac{1}{\hat{\sigma}_X} \xrightarrow{P} g(1) = 1 \tag{5}$$

が分かる. $1/\hat{\sigma}_Y$ についても同様にして,

$$\frac{1}{\hat{\sigma_Y}} \xrightarrow{P} 1 \tag{6}$$

を得る. よって (5), (6) より

$$\frac{1}{\hat{\sigma}_X\hat{\sigma}_Y} \xrightarrow{P} 1 \tag{7}$$

が成り立つ.

また, $g'(a_n)$ と $g'(b_n)$ の確率収束先も調べる. $g'(x)$ は連続関数であるから

$$\hat{\sigma}_X^2 \xrightarrow{P} 1 \Rightarrow a_n \xrightarrow{P} 1 \Rightarrow g'(a_n) \xrightarrow{P} g'(1) = -\frac{1}{2} \tag{8}$$

が分かり, $g'(b_n)$ についても同様に

$$g'(b_n) \xrightarrow{P} -\frac{1}{2} \tag{9}$$

が成り立つ.

これまでの結果を元に, 0 に確率収束する $o_p(1)$ を使って, (1) 右辺の各項を変形する.

$\rho\sqrt{n}\,(1/\hat{\sigma}_X - 1)$ については, (2),(3),(8) より,

$$\rho\sqrt{n}\left(\frac{1}{\hat{\sigma}_X} - 1\right) = -\frac{\rho}{2}\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\left(Z_{1i}^2 - 1\right) + o_p(1) \tag{10}$$

と表せられる.

$\rho/\hat{\sigma}_X \cdot \sqrt{n}\,(1/\hat{\sigma}_Y - 1)$ についても, (2),(4),(5),(9) より,

$$\begin{aligned}&\frac{\rho}{\hat{\sigma}_X}\sqrt{n}\left(\frac{1}{\hat{\sigma}_Y} - 1\right) \\ &= -\frac{\rho}{2}\sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\Big\{(\rho Z_{1i} + \sqrt{1-\rho^2}Z_{2i})^2 - 1\Big\} + o_p(1)\end{aligned} \tag{11}$$

とできる.

$1/(\hat{\sigma}_X\hat{\sigma}_Y) \cdot \sqrt{n}(\hat{\sigma}_{XY} - \rho)$ については, (2),(7) より,

$$\begin{aligned}&\frac{1}{\hat{\sigma}_X\hat{\sigma}_Y}\sqrt{n}(\hat{\sigma}_{XY} - \rho) \\ &= \sqrt{n}\Big\{\frac{1}{n}\sum_{i=1}^{n} Z_{1i}(\rho Z_{1i} + \sqrt{1-\rho^2}Z_{2i}) - \rho\Big\} + o_p(1)\end{aligned} \tag{12}$$

となる. よって, (10),(11),(12) より $\sqrt{n}(\hat{\rho} - \rho)$ は,

$$\begin{aligned}\sqrt{n}(\hat{\rho} - \rho) = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n}\Big\{& Z_{1i}(\rho Z_{1i} + \sqrt{1-\rho^2}Z_{2i}) \\ &- \frac{\rho}{2}(\rho Z_{1i} + \sqrt{1-\rho^2}Z_{2i})^2 - \frac{\rho}{2}Z_{1i}^2\Big\} + o_p(1)\end{aligned} \tag{13}$$

と変形できる. ここで $i = 1, \cdots, n$ に対して,

$$\begin{aligned}V_i \equiv\ & Z_{1i}(\rho Z_{1i} + \sqrt{1-\rho^2}Z_{2i}) \\ &- \frac{\rho}{2}(\rho Z_{1i} + \sqrt{1-\rho^2}Z_{2i})^2 - \frac{\rho}{2}Z_{1i}^2\end{aligned}$$

と置けば, (13) は

$$\sqrt{n}(\hat{\rho} - \rho) = \sqrt{n}\frac{1}{n}\sum_{i=1}^{n} V_i + o_p(1) \tag{14}$$

と表せる. V_i を変形すると,

$$V_i = (1-\rho^2)\left(\frac{\rho}{2}Z_{1i}^2 + \sqrt{1-\rho^2}Z_{1i}Z_{2i} - \frac{\rho}{2}Z_{2i}^2\right)$$

となり, その期待値と分散を求めると,

尖度 $\mathrm{E}(Z_{1i}^4) - 3 = \mathrm{E}(Z_{2i}^4) - 3 = 0$ と,

Z_{1i} と Z_{2i} は互いに独立であることを利用して,

$$\begin{aligned}\mathrm{E}(V_i) &= (1-\rho^2)\mathrm{E}\left(\frac{\rho}{2}Z_{1i}^2 + \sqrt{1-\rho^2}Z_{1i}Z_{2i} - \frac{\rho}{2}Z_{2i}^2\right) \\ &= (1-\rho^2)\left(\frac{\rho}{2} - \frac{\rho}{2}\right) = 0, \\ \mathrm{V}(V_i) &= \mathrm{E}\left[(1-\rho^2)^2\left(\frac{\rho}{2}Z_{1i}^2 + \sqrt{1-\rho^2}Z_{1i}Z_{2i} - \frac{\rho}{2}Z_{2i}^2\right)^2\right] \\ &= (1-\rho^2)^2\end{aligned}$$

となる. よって中心極限定理から,

$$\sqrt{n}\frac{1}{n}\sum_{i=1}^{n} V_i \xrightarrow{\mathcal{L}} \mathrm{N}\left(0, (1-\rho^2)^2\right)$$

となり, (14) からスラツキーの定理を適用して結論を得る.

□

定理 1(フィッシャーの z 変換)

補題 1 の仮定において,

$$Z(r) \equiv \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)(r \in (-1,1))$$

としたとき, $n \to \infty$ として

$$\sqrt{n-3}\{Z(\hat{\rho}) - Z(\rho)\} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$$

が成り立つ.

証明

$$Z'(r) = \frac{1}{2}\left(\frac{1}{1+r} + \frac{1}{1-r}\right) = \frac{1}{1-r^2}$$

より, 補題 1 とデルタ法から,
確率変数 $W \sim \mathrm{N}\left(0, (1-\rho^2)^2\right)$ を用いて,

$$\begin{aligned}\sqrt{n-3}\{Z(\hat{\rho}) - Z(\rho)\} &= \frac{\sqrt{n-3}}{\sqrt{n}}\sqrt{n}\{Z(\hat{\rho}) - Z(\rho)\} \\ \xrightarrow{\mathcal{L}} \lim_{n\to\infty}\frac{\sqrt{n-3}}{\sqrt{n}}\frac{1}{1-\rho^2}W &= \frac{1}{1-\rho^2}W \sim \mathrm{N}(0,1)\end{aligned}$$

となるので, 題意を満たす. □

4 漸近的な検定方式と同時信頼区間

定理 1 で定義した関数 $Z(r)$ の性質を調べる.
$Z(r) = \tanh^{-1} r$ であり,

$$\begin{aligned}\frac{d}{d\theta}Z^{-1}(\theta) = \frac{d}{d\theta}\tanh\theta &= \frac{(e^\theta + e^{-\theta})^2 - (e^\theta - e^{-\theta})^2}{(e^\theta + e^{-\theta})^2} \\ &= 1 - \tanh^2\theta > 0\end{aligned}$$

から, Z^{-1} は狭義単調増加であるので,

$$\theta_1 < \theta_2 \Leftrightarrow Z^{-1}(\theta_1) < Z^{-1}(\theta_2) \Leftrightarrow \tanh\theta_1 < \tanh\theta_2 \tag{15}$$

が成り立つ. ここで, 第 i 群 $(i = 1, \cdots, k)$ にて,

$$\hat{\rho}_i \equiv \frac{1/n_i \cdot \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i.)(Y_{ij} - \bar{Y}_i.)}{\sqrt{1/n_i \cdot \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i.)^2}\sqrt{1/n_i \cdot \sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_i.)^2}}$$

と置いたとき, 定理 1 から, $n_i \to \infty$ として,

$$\sqrt{n_i - 3}\{Z(\hat{\rho}_i) - Z(\rho_i)\} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1) \tag{16}$$

が成り立つ. このことを利用して, 漸近的な検定方式と信頼区間を求める.

4.1 多重比較検定

帰無仮説 H_{0i} と対立仮説 H_{0i}^{A*} の漸近的な多重比較検定について述べる. $T_i(r) \equiv \sqrt{n_i - 3}\{Z(\hat{\rho}_i) - Z(r)\}(i = 1, \cdots, k)$ とし, $\mathrm{N}(0,1)$ における上側 $100 \times h(\alpha)\%$ 点を $z_{h(\alpha)}$ とする. $\min N$ を集合 N の元の最小値として,

$$\lim_{\min N\to\infty} \mathrm{P}\left(\max_{1\le i\le k}|T_i(\rho_i)| \le z_{h_1(\alpha)}\right) = 1 - \alpha\ , \tag{17}$$

$$\lim_{\min N\to\infty} \mathrm{P}\left(\max_{1\le i\le k}T_i(\rho_i) \le z_{h_2(\alpha)}\right) = 1 - \alpha \tag{18}$$

が成り立つような $h_1(\alpha)$, $h_2(\alpha)$ をそれぞれ考える. $T_1(\rho_1), \cdots, T_k(\rho_k)$ は互いに独立より,

$$\mathrm{P}\left(\max_{1\le i\le k}|T_i(\rho_i)| \le z_{h_1(\alpha)}\right) = \prod_{i=1}^{k}\mathrm{P}\left(|T_i(\rho_i)| \le z_{h_1(\alpha)}\right)\ ,$$

$$\mathrm{P}\left(\max_{1\le i\le k}T_i(\rho_i) \le z_{h_2(\alpha)}\right) = \prod_{i=1}^{k}\mathrm{P}\left(T_i(\rho_i) \le z_{h_2(\alpha)}\right)$$

が成り立つ. よって (16) から,

$$\begin{aligned}(17) &\Leftrightarrow \lim_{n_i\to\infty}\mathrm{P}\left(|T_i(\rho_i)| \le z_{h_1(\alpha)}\right) = (1-\alpha)^{1/k} \\ &\Leftrightarrow \lim_{n_i\to\infty}\mathrm{P}\left(T_i(\rho_i) \le z_{h_1(\alpha)}\right) = \frac{1 + (1-\alpha)^{1/k}}{2} \\ &\Leftrightarrow h_1(\alpha) = \lim_{n_i\to\infty}\mathrm{P}\left(T_i(\rho_i) \ge z_{h_1(\alpha)}\right) \\ &\qquad = 1 - \frac{1 + (1-\alpha)^{1/k}}{2} = \frac{1 - (1-\alpha)^{1/k}}{2}\ , \\ (18) &\Leftrightarrow \lim_{n_i\to\infty}\mathrm{P}\left(T_i(\rho_i) \le z_{h_2(\alpha)}\right) = (1-\alpha)^{1/k} \\ &\Leftrightarrow h_2(\alpha) = \lim_{n_i\to\infty}\mathrm{P}\left(T_i(\rho_i) \ge z_{h_2(\alpha)}\right) \\ &\qquad = 1 - (1-\alpha)^{1/k}\end{aligned}$$

となり, $h_1(\alpha)$ と $h_2(\alpha)$ の式が導かれる.

以上から, $\{$ 帰無仮説 $H_{0i} \mid 1 \le i \le k\}$ に対する水準 α の漸近的な多重比較検定として, $i = 1, \cdots, k$ に対して次の ①,②,③ が与えられる.

① 両側検定：帰無仮説 H_{0i} vs. 対立仮説 $H_{0i}^{A\pm}$ のとき, $|T_i(\rho_{0i})| > z_{h_1(\alpha)}$ となる i に対して帰無仮説 H_{0i} を棄却し, 対立仮説 $H_{0i}^{A\pm}$ を受け入れ, $\rho_i \ne \rho_{0i}$ と判定する.

② 片側検定：帰無仮説 H_{0i} vs. 対立仮説 H_{0i}^{A+} のとき, $T_i(\rho_{0i}) > z_{h_2(\alpha)}$ となる i に対して帰無仮説 H_{0i} を棄却し, 対立仮説 H_{0i}^{A+} を受け入れ, $\rho_i > \rho_{0i}$ と判定する.

③ 片側検定：帰無仮説 H_{0i} vs. 対立仮説 H_{0i}^{A-} のとき, $T_i(\rho_{0i}) < -z_{h_2(\alpha)}$ となる i に対して帰無仮説 H_{0i} を棄却し, 対立仮説 H_{0i}^{A-} を受け入れ, $\rho_i < \rho_{0i}$ と判定する.

4.2 同時信頼区間

相関係数の漸近的な同時信頼区間を求める.
$|T_i(\rho_i)| \le z_{h_1(\alpha)}$, $T_i(\rho_i) < z_{h_2(\alpha)}$, $T_i(\rho_i) > -z_{h_2(\alpha)}$ の 3 式をそれぞれ ρ_i について解くと, (15) から,

$$\begin{aligned}&|T_i(\rho_i)| \le z_{h_1(\alpha)} \\ &\Leftrightarrow -z_{h_1(\alpha)} \le \sqrt{n_i - 3}\{Z(\hat{\rho}_i) - Z(\rho_i)\} \le z_{h_1(\alpha)} \\ &\Leftrightarrow \tanh\left(Z(\hat{\rho}_i) - \frac{z_{h_1(\alpha)}}{\sqrt{n_i - 3}}\right) \\ &\qquad \le \rho_i \le \tanh\left(Z(\hat{\rho}_i) + \frac{z_{h_1(\alpha)}}{\sqrt{n_i - 3}}\right)\ ,\end{aligned}$$

$$\begin{aligned}&T_i(\rho_i) < z_{h_2(\alpha)} \Leftrightarrow \sqrt{n_i - 3}\{Z(\hat{\rho}_i) - Z(\rho_i)\} < z_{h_2(\alpha)} \\ &\Leftrightarrow \tanh\left(z(\hat{\rho}_i) - \frac{\zeta_{h_2(\alpha)}}{\sqrt{n_i - 3}}\right) < \rho_i\ ,\end{aligned}$$

$$\begin{aligned}&T_i(\rho_i) > -\zeta_{f_2(\alpha)} \Leftrightarrow \sqrt{n_i - 3}\{Z(\hat{\rho}_i) - Z(\rho_i)\} > -z_{h_2(\alpha)} \\ &\Leftrightarrow \rho_i < \tanh\left(Z(\hat{\rho}_i) + \frac{z_{h_2(\alpha)}}{\sqrt{n_i - 3}}\right)\end{aligned}$$

となる．よって，ρ_i $(i=1,\cdots,k)$ についての信頼係数 $1-\alpha$ の漸近的な同時信頼区間は，次の ①,②,③ で与えられる．

① 両側信頼区間：

$$\tanh\left(Z(\hat{\rho}_i)-\frac{z_{h_1(\alpha)}}{\sqrt{n_i-3}}\right)\leq\rho_i$$
$$\leq\tanh\left(Z(\hat{\rho}_i)+\frac{z_{h_1(\alpha)}}{\sqrt{n_i-3}}\right)\quad(i=1,\cdots,k).$$

② 上側片側信頼区間：

$$\tanh\left(Z(\hat{\rho}_i)-\frac{z_{h_2(\alpha)}}{\sqrt{n_i-3}}\right)<\rho_i\ (i=1,\cdots,k)\ .$$

③ 下側片側信頼区間：

$$\rho_i<\tanh\left(Z(\hat{\rho}_i)+\frac{z_{h_2(\alpha)}}{\sqrt{n_i-3}}\right)\ (i=1,\cdots,k)\ .$$

5 c 言語プログラムの解説とデータ解析

c 言語により，上で求めた検定方式と信頼区間を元にデータ解析を行うプログラムを作成した．

5.1 main プログラムの流れ

1. MODEL 関数で，データ情報を入力しモデルを作成．
2. MEAN 関数で，myuhX, myuhY に標本平均を格納．
3. VAR_COV_CORR 関数で，sigmahX, sigmahY に標本分散，Covh に標本共分散，rhoh に標本相関係数を格納．
4. HYOHONTI 関数で，それぞれの群の標本値を表示．
5. KENTEI 関数で，検定の実行．
6. SHINRAIKUKAN 関数で，信頼区間の表示．

5.2 データ内容

脊椎動物における各種属の寿命と全長から，相関係数を解析する．X 標本を寿命 [年]，Y 標本を全長 [cm] とみなす．

表 2 各種族の寿命と全長のデータ

群	種族名	サイズ	データ
第 1 群	爬虫類	121	$(75,90),\cdots,(75,60)$
第 2 群	両生類	46	$(10,4.5),\cdots,(12.5,13.5)$
第 3 群	哺乳類	297	$(71,166.5),\cdots,(25,25)$
第 4 群	鳥類	45	$(72,120),\cdots,(10,17)$

5.3 実行結果

水準 α を 0.05, 比較のための相関係数を全て 0 としたとき，両側検定における実行結果及び信頼区間は以下の通りとなった．

```
(1) 両側検定:
帰無仮説 H_{0,1}を棄却し，対立仮説 H_{0,1}^{A ±}を
受け入れ， ρ_1 ≠ρ_{0,1}と判定.
帰無仮説 H_{0,2}を棄却し，対立仮説 H_{0,2}^{A ±}を
受け入れ， ρ_2 ≠ρ_{0,2}と判定.
帰無仮説 H_{0,3}を棄却し，対立仮説 H_{0,3}^{A ±}を
受け入れ， ρ_3 ≠ρ_{0,3}と判定.
帰無仮説 H_{0,4}を棄却し，対立仮説 H_{0,4}^{A ±}を
受け入れ， ρ_4 ≠ρ_{0,4}と判定.
(1) 両側信頼区間:
0.01 ≦ρ_1 ≦ 0.43(i=1).
0.43 ≦ρ_2 ≦ 0.84(i=2).
0.62 ≦ρ_3 ≦ 0.77(i=3).
0.11 ≦ρ_4 ≦ 0.71(i=4).
```

5.4 考察

帰無仮説はすべて棄却されているため，どの種族も寿命と全長には相関があると考えられる．特に哺乳類に関しては強い正の相関がある．哺乳類の標本には人間も含めているが，人間に比べてクジラなど全長が何メートルに及ぶ生き物でさえ 60 年も生きないことと，この強い正の相関が見られることは，人間が全長に対して寿命が異常にあることを示していると推測する．爬虫類，両生類，鳥類に関して両側信頼区間の幅が広いのは，亀などの全長が小さくて寿命が長い生き物であったり，ペリカンなどの全長が大きくて寿命が短い生き物が一定数存在することによると考えられる．

6 おわりに

本論文では，多群 2 次元正規分布モデルにおける相関係数の多重比較検定法について考察した．標本相関係数の分布収束性について，予想以上に計算を要した．フィッシャーの z 変換から，標準正規分布に従わせることで，同時信頼区間を求めることができた．正規分布モデルを仮定しない場合，4 次のモーメントが計算できないことから，今回のアプローチでは相関係数の分布収束性を示すことが不可能であると考える．

参考文献

[1] 白石高章：『多群連続モデルにおける多重比較法―パラメトリック，ノンパラメトリックの数理統計』共立出版，東京，2018

[2] 白石高章：『統計科学の基礎―データと確率の結びつきがよくわかる数理』日本評論社，東京，2019

[3] 清水邦夫：『相関係数』近代科学社，東京，2020