(20) # 級内相関係数の標本分布について

鍋谷　清治

級内相関係数の標本分布は最初 R.A.Fisher によつて求められたのであるが，これに対して当講究録第7巻第5号に於て小川先生は解析的な証明を與えられた。

Fisher はその著書に於て級内相関を取扱う際に，1つの級に属する観測値 $x_1, x_2, \cdots\cdots, x_k$ は，級による変動を表わす値 y と，random な変動を表わす値 $z_1, z_2, \cdots\cdots z_k$ との和，$y+z_1, y+z_2, \cdots\cdots, y+z_k$ として表わされるものと考えている。

この場合，y の分散を A，z の分散を B として y, z_1, z_2, z_k は互に独立に分布するものとすれば，1つの観測値 $x = y+z$ の分散は $A+B$，同一の級に属する異なる観測値（例えば $x_1 = y+z_1$ と $x_2 = y+z_2$）の間の共分散は A となるから，その間の相関係数は $\rho = \frac{A}{A+B}$ となる。

これを母集団に於ける級内相関の値として，これに対する推定値 r の分布を分散分析の方法で求めている。

勿論 y も z も正規分布に従うと仮定しての上である。しかしこの方法で取扱えるのは $\rho \geqq 0$ の場合（1つの級に属するもの個数に制限がなければ必ず $\rho \geqq 0$ となる）だけあつて，$\rho < 0$ の場合にまでその考えを拡張するのは不自然のように思われる。

これに対して小川先生のやり方では $\rho \geqq 0$ の制限が要らないけれども，分布の導き方が余りにも面倒なので，私はここで，$Fisher$ の方法を少しく改良した形で ρ の符号には無関係な導き方を述べることにする。

なお，その証明方法については，遠藤氏との共訳

フィッシャー　研究者のための 統計的方法

（原著　R.A.Fisher : Statistical method for research works）

の補註にもその概要を述べてある。

上述のように1つの級に属する観測値を $x_1, x_2, ----, x_k$ とする。そしてこれ等 k 個の量はすべて同一平均 μ，同一分散 σ^2 を有する k 変量の正規分布に従い，しかも異なる観測値の間の相関は一定量 ρ に等しいとする。このρが母相関であつて，

$-\frac{1}{k-1} < \rho < 1$ とする。

標本の大きさを n（従つて kn 個の観測値がある）とし，1つの級に於ける k 個の観測値の平均を $\bar{x}$，kn 個全体の観測値の平均を $\bar{\bar{x}}$ とすれば，

$$\overset{kn}{\underset{1}{S}}(x-\bar{\bar{x}})^2 = k\overset{n}{\underset{1}{S}}(\bar{x}-\bar{\bar{x}})^2 + \overset{kn}{\underset{1}{S}}(x-\bar{x})^2$$

であつて，標本級内相関係数 r は

$$(1)\qquad \frac{k\overset{n}{\underset{1}{S}}(\bar{x}-\bar{\bar{x}})^2}{\overset{kn}{\underset{1}{S}}(x-\bar{x})^2} = \frac{1+(k-1)r}{(k-1)(1-r)}$$

から求まる。

上述の仮定があれば，$x_1, x_2, - \cdots, x_k$ の分散行列は

$$V = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \cdots & \cdots & \cdots & \cdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

となり，$x_1, x_2, \cdots\cdots, x_k$ から出来る列－vector を $\mathfrak{x}$．

$$A = \begin{pmatrix} 1-\frac{1}{k} & -\frac{1}{k} & \cdots\cdots & -\frac{1}{k} \\ -\frac{1}{k} & 1-\frac{1}{k} & \cdots\cdots & -\frac{1}{k} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ -\frac{1}{k} & -\frac{1}{k} & \cdots\cdots & 1-\frac{1}{k} \end{pmatrix}, \quad \alpha = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \\ \vdots \\ \frac{1}{k} \end{pmatrix}$$

と置けば，

$$\bar{x} = \alpha' \mathfrak{x}, \quad \mathop{S}_1^k (x_i - \bar{x})^2 = \mathfrak{x}' A \mathfrak{x}$$

となる。 ここで行列の計算から

$$A V \alpha = 0$$

によつて $\bar{x}$ と $\mathfrak{x}' A \mathfrak{x}$ との間の独立性がわかるし，

$$A V A = \sigma^2 (1-\rho) A$$

によつて，$\mathfrak{x}' A \mathfrak{x}$ は $\sigma^2(1-\rho)$ で割つたものは χ^2 分布に従うことがわかる。

更に A の階数は $k-1$ で

$$|V| = \sigma^{2k} (1-\rho)^{k-1} \{ 1 + (k-1)\rho \} \neq 0$$

であるから，その χ^2 分布の自由度は $k-1$ 個になる。

随つて，(1) の分母を $\sigma^2(1-\rho)$ で割れば，$(k-1)n$ 個の自由度を持つ χ^2 分布に従う。

他方 $\bar{x}$ の分散は

$$\alpha' V \alpha = \sigma^2 \frac{1+(k-1)\rho}{k}$$

であるから，$\mathop{S}_1^n (\bar{x} - \bar{\bar{x}})^2$ をこれで割つたものは，自由度が，$n-1$ 個の χ^2 分布に従う。

しかもその分布は上述のように $\overset{kn}{\underset{1}{S}}(x-\bar{x})^2$ とは独立である。
随って

$$F = \frac{k\overset{n}{\underset{1}{S}}(\bar{x}-\bar{\bar{x}})^2}{\sigma^2\{1+(k-1)\rho\}(n-1)} \Big/ \frac{\overset{kn}{\underset{1}{S}}(x-\bar{x})^2}{\sigma^2(1-\rho)(k-1)n}$$

$$(2) \qquad = \frac{1+(k-1)r}{1-r} \cdot \frac{1-\rho}{1+(k-1)\rho} \cdot \frac{n}{n-1}$$

は分子の自由度が $n-1$，分母の自由度が $(k-1)n$ の F 分布

$$\frac{\Gamma\left(\frac{kn-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{(k-1)n}{2}\right)}(n-1)^{\frac{n-1}{2}}\{(k-1)n\}^{\frac{(k-1)n}{2}} \frac{F^{\frac{n-3}{2}}}{\{(k-1)n+(n-1)F\}^{\frac{kn-1}{2}}}dF$$

に従う。

これに対して（2）の変換を行えば r の分布

$$\frac{\Gamma\left(\frac{kn-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{(k-1)n}{2}\right)} k^{-\frac{kn-3}{2}}(k-1)^{\frac{(k-1)n}{2}}(1-\rho)^{\frac{n-1}{2}}\{1+(k-1)\rho\}^{\frac{(k-1)n}{2}}$$

$$\times(1-r)^{\frac{(k-1)n-2}{2}}\{1+(k-1)r\}^{\frac{n-3}{2}}\{1+(k-2)\rho-(k-1)\rho r\}^{-\frac{kn-1}{2}}dr$$

が得られる。

参 照 文 献

R. A. Fisher : On the "probable error" of a coefficient of correlation deduced from a small sample. Metron 1 (1921)

小 川 潤 次 郎 ： 級内相関係数の標本分布の解析的な導出について。 講究録，第7巻第5号

受付 1951. 12. 19.